Cuk - Ho 1636 - 69 - Pb 1491S
1-4
69

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

REINHOLD BAER
UNIVERSITY OF ILLINOIS

WEI-LIANG CHOW
THE JOHNS HOPKINS UNIVERSITY

ANDRÉ WEIL
UNIVERSITY OF CHICAGO

AUREL WINTNER
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

S. S. CHERN
C. CHEVALLEY
J. A. DIEUDONNÉ
A. M. GLEASON
HARISH-CHANDRA
P. HARTMAN
G. P. HOCHSCHILD
I. KAPLANSKY
E. R. KOLCHIN
W. S. MASSEY
D. C. SPENCER
A. D. WALLACE

PUBLISHED UNDER THE JOINT AUSPICES OF
THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

510.5
.002

VOLUME LXXVII
1955

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

THE JOHNS HOPKINS PRESS
BALTIMORE 18, MARYLAND
U. S. A.

P16915

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

INDEX

PAGE

Abhyankar, Shreeram. On the ramification of algebraic functions, . 575

Auslander, Maurice and R. C. Lyndon. Commutator subgroups of free groups, ... 929

Bass, Robert W. On the regular solutions at a point of singularity of a system of non-linear differential equations, ... 734

Boas, R. P., Jr. Isomorphism between H^p and L^p, ... 655

Chevalley, Claude. Invariants of finite groups generated by reflections, ... 778

Dieudonné, Jean. Lie groups and Lie hyperalgebras over a field of characteristic $p > 0$ (II), ... 218

Dieudonné, Jean. Lie groups and Lie hyperalgebras over a field of characteristic $p > 0$ (IV), ... 429

Drazin, M. P. Engel rings and a result of Herstein and Kaplansky, . 895

Ehrenpreis, Leon. Solution of some problems of division, . . . 293

Ehrenpreis, Leon. Appendix to the paper "Mean periodic functions, I," 731

Goffman, Casper. Convergence in area of integral means, . . . 563

Hamstrom, Mary-Elizabeth. A characterization of plane continuous curves which are filled up by continuous collections of continuous curves, ... 914

Harish-Chandra. Representations of semisimple Lie groups IV, . 743

Hartman, Philip and Aurel Wintner. Asymptotic integrations of linear differential equations, ... 45

Errata, ... 932

Hartman, Philip and Aurel Wintner. On uniform Dini conditions in the theory of linear partial differential equations of elliptic type, ... 329

Hartman, Philip and Aurel Wintner. On the local behavior of solutions of non-parabolic partial differential equations (III), 453

Hartman, Philip and Aurel Wintner. On the assignment of asymptotic values for the solutions of linear differential equations of second order, ... 475

GENERALIZED GALOIS THEORY FOR RINGS WITH MINIMUM CONDITION, II.*

By Tadasi Nakayama.

In our first paper, "Generalized Galois theory for rings with minimum condition, *Amer. J. Math.*, vol. 73 (1951), pp. 1-12—referred to as I—we gave a certain Galois theory for primary rings with minimum condition. The theory contains the Cartan-Jacobson [1], [2] Galois theory for sfields, at least as far as the main Galois correspondence is concerned. But the "inner" portion of the Galois group, there considered, comes from a certain completely primary subring. Now that we are dealing with a general primary ring, satisfying the minimum condition, it is desirable to consider a wider class of Galois groups whose inner portions come from general primary subrings rather than from completely primary ones. In fact in another paper [5] we have developed a Galois theory for simple rings, with minimum condition, in which the Galois groups have rather satisfactory generality and have inner portions consisting of inner automorphisms effected by regular elements of simple subalgebras. Combining the methods and results of this paper with those in I, we now give a Galois theory for primary rings which deals with Galois groups of the generality required above, though it fails to cover the theory of I completely. Our main Galois correspondence is given in **2**, Theorem 1, our "regular" Galois group being defined both at the opening of **2** and in succession to Theorem 1.

Moreover, after a study of a certain special type of factor-groups of the Galois group (**3**, Theorem 2), we prove an extension theorem for isomorphisms (**3**, Theorem 3), to recover a second main aspect of Galois theory which was neglected in our first paper I.

(*Corrections*: The condition i) in I, Theorem 2 was stated incorrectly. There we were to assume further that K has an independent Z-basis such that the corresponding left-multiplications on R constitute an independent R_r-basis of R_rK_l. On the other hand, the condition iv) there was stated in a rather awkward form. We had to assume only that every inner automorphism of R effected by a regular element of K belongs to Φ. (For the fact that this, supported by the other conditions, implies iv), cf. **1** below.)

---

* Received June 30, 1954.

Further, I, Lemma 4 should, as G. Azumaya has kindly pointed out, refer only to a primary ring R; the primarity is needed to secure the existence of a system of elements $a_1, a_2, \cdots, a_s$ in K, asserted on page 7, line 17 there, such that $\gamma_1, \cdots, \gamma_t, a_{1l}, \cdots, a_{sl} \bmod N_rK_l$ together form an independent R_r/N_r-right-basis of R_rK_l/N_rK_l. But, since we considered only primitive rings in Theorem 2 there, the validity of our main result of the paper I remains unaffected.

The same lemma was used also in an older paper [3]. The writer wishes to take this opportunity to make the correction that the final theorem (3. 2) in that paper should refer only either to a primary ring R or to the case where K modulo the radical N of R contains the center of the residue-ring R/N; it is easy to see that also in the last case the existence of $a_1, \cdots, a_s$ as above is secured.)

1. Preliminaries. We first cite some of our previous results which we need in the sequel. Let R be a ring satisfying the minimum condition (for left- and right-ideals) and possessing a unit element 1. With a subset X of R we denote by X_l, X_r the sets of the left and the right multiplications in R by the elements of X. Let $\mathfrak{A}$ be the absolute endomorphism ring of R as a module. If $\mathfrak{F}$ is any subset of $\mathfrak{A}$, the commuter subring $V_{\mathfrak{A}}(\mathfrak{F})$ of $\mathfrak{F}$ in $\mathfrak{A}$ is nothing but the $\mathfrak{F}$-endomorphism ring of R. In particular, if Φ ($\subseteq \mathfrak{A}$) is a group of automorphisms of R and if S is the invariant system of Φ in R, then $S_r = V_{\mathfrak{A}}(R_l\Phi)$. First, we consider an automorphism group Φ of R satisfying the following conditions:

i_0) Φ is complete in the sense that the subgroup Φ_0 of Φ consisting of the inner automorphisms of R contained in Φ is the set of all inner automorphisms effected by the regular elements of a certain subring K of R containing the center Z of R,

i_1) K is generated by its regular elements,

i_2) K has a (linearly) independent finite basis $a_1 = 1, a_2, \cdots, a_k$ over Z such that $a_{1r}, a_{2r}, \cdots, a_{kr}$ are independent over R_l,

i_3) R is a regular module of R_lK_r, or, which amounts to the same, the direct sum R^k of k copies of R is R_lK_r-right-isomorphic with R_lK_r itself (For the notions of regular modules and their ranks cf. I, footnote 7),

ii) The factor group $G = \Phi/\Phi_0$ is finite, and if $\{\rho(1) = 1, \rho(\sigma), \cdots, \rho(\tau)\}$ is a representative system of Φ modulo Φ_0 $(G = \{1, \sigma, \cdots, \tau\})$, the R_l-two-sided modules R_l, $\rho(\sigma)R_l$ $(\sigma \neq 1)$ have no isomorphic composition residue-modules (If this holds for any one representative system then the same is the case for any other representative system).

If Φ satisfies these conditions, then R is regular also with respect to the operator ring $R_l\Phi$ and we have not only $S_r = V_{\mathfrak{A}}(R_l\Phi)$ but also

$$V_{\mathfrak{A}}(S_r) = R_l\Phi = \sum_{\sigma} \rho(\sigma)R_lK_r, \tag{1}$$

where S is the invariant system of Φ in R. Moreover, R possesses an independent right-basis of $(G:1)k$ terms over S (I, Theorem 1; everything is stated here in the left-right symmetric form to I). The number $(G:1)k = (\Phi:\Phi_0)k$ is called the *reduced order* of Φ.

LEMMA 0. *Let R be primary and let its subring K satisfy the condition* i_2). *Let $\mathfrak{M}$ be an R_l-two-sided submodule of R_lK_r such that $\mathfrak{M}N_l = \mathfrak{M} \cap N_lK_r$, where N is the radical of R, and $\mathfrak{M}/\mathfrak{M}N_l$ is R_l/N_l-right-regular* (*These last two conditions are certainly satisfied when $\mathfrak{M}$ is R_l-right-regular, because of the primarity of R*). *Then $\mathfrak{M}$ is a direct summand of the R_l-two-sided module R_lK_r and has also an independent finite basis over R_l consisting of elements of K_r.*

Our old proof, in I, Lemma 4, remains valid, since the existence of $a_1, a_2, \cdots, a_s$ as asserted there is now true; cf. the correction in the introduction.

Let R be primary and let ϕ be an automorphism of R leaving the invariant system S of Φ elementwise fixed, Φ being as above. Then $\phi \in V_{\mathfrak{A}}(S_r) = \sum_{\sigma} \rho(\sigma)R_lK_r$ and $\phi R_l = R_l\phi$. So ϕR_l $(= R_l\phi)$ is an R_l-two-sided submodule of $\sum \rho(\sigma)R_lK_r$, and is, as such, a direct sum of submodules of $\rho(\sigma)R_lK_r$, σ running over G; observe the condition ii). As R (whence R_l) is primary (and is thus two-sided directly indecomposable), we see that there exists a σ such that (cf. I, Lemma 2)

$$\phi \in \rho(\sigma)R_lK_r. \tag{2}$$

Putting $\phi_0 = \rho(\sigma)^{-1}\phi$, we have $\phi_0 \in R_lK_r$. Also ϕ_0 is an automorphism of R, leaving S elementwise fixed, and thus $\phi_0R_l = R_l\phi_0$. It follows from Lemma 0 that ϕ_0R_l $(= R_l\phi_0)$ has the form R_la_r $(a \in K)$. Putting $\phi_0 = b_la_r$ $(b \in R)$ and observing $1^{\phi_0} = 1$ we see that (a is regular and) $b = a^{-1}$. Thus ϕ_0 is an inner automorphism of R effected by a regular element of K. So $\phi_0 \in \Phi_0$ and $\phi = \rho(\sigma)\phi_0 \in \Phi$; this point was left rather obscure in I. Hence, Φ *exhausts the automorphisms of R leaving S elementwise fixed* (*provided that R is primary*).

Another use we want to make of Lemma 0 is, as in I, to study the structure of the endomorphism ring of R with respect to the right multiplication of a certain subring of R. Let T be a subring of R which contains the invariant system S of our automorphism group Φ (satisfying the above conditions). Then the T_r-endomorphism ring of R is a subring of $V_{\mathfrak{A}}(S_r) = \sum \rho(\sigma) R_l K_r$ and contains R_l. By virtue of ii) we have

$$V_{\mathfrak{A}}(T_r) = \sum \rho(\sigma) U_\sigma \tag{3}$$

with certain R_l-two-sided submodules U_σ of $R_l K_r$. Suppose that R is T_r-regular. Then $V_{\mathfrak{A}}(T_r)$ is R_l-right-regular. Hence each U_σ is a direct sum of submodules R_l-right-isomorphic to directly indecomposable right-ideal direct components of R_l. R being assumed to be primary, this means that each U_σ is R_l-right-regular. So our lemma may be applied to yield

LEMMA 1. *Let R be primary and T_r-regular, T being a subring of R containing S. Then each U_σ in* (3) *has the form $U_\sigma = R_l L_r^{(\sigma)}$, where $L^{(\sigma)}$ is a subring of K containing Z and possessing an independent finite basis over Z such that the corresponding right multiplications on R constitute an independent basis of U_σ over R_l. We have, therefore,*

$$V_{\mathfrak{A}}(T_r) = \sum_{\sigma} \rho(\sigma) R_l L_r^{(\sigma)}. \tag{4}$$

Now, the case of a simple ring R was studied in our paper [5]. We need the following main result there:

Let R be a simple ring, satisfying minimum condition and possessing unit element 1, and let Φ be an automorphism group of R which is regular in the sense that it satisfies i_0) with a simple subring K finite over the center Z of R and Φ/Φ_0 is finite; then i_1), i_2), i_3) and ii) are all automatically fulfilled. The invariant system S of Φ in R is a simple subring of R whose commuter $V_R(S)$ in R is K. The regular subgroups of Φ are in 1-1 Galois correspondence with the simple subrings of R containing S and possessing simple commuters in R, and any isomorphism between two such simple subrings of R leaving S elementwise fixed is induced by an element of Φ ([5], Theorems 5, 6).

2. Galois theory for primary rings. Let R be a primary ring, with minimum condition and unit element 1. Let N be its radical, and Z its center. We call a group of automorphisms of R *regular* when it satisfies the conditions i_0), i_2) and ii) in **1** and when, moreover, the subring K is such that $((K+N)/N$ is simple and) the product in R/N of $(K+N)/N$ and the center of R/N is simple. It is clear that this definition is in accord with the notion of a regular automorphism group for the case of a simple ring R, as defined in [5] (cf. the end of the preceding section).

For any subset A of $\mathfrak{A}$ leaving N set-wise invariant we denote by A^0 the set of endomorphisms of the residue-module R/N induced by the elements of A; we identify equal endomorphisms of R/N induced by distinct elements of A. So, in particular, R_l^0 and R_r^0 are the left- and right-multiplication rings of the residue ring R/N.

Now we observe that a regular automorphism group, as just defined, satisfies also the conditions i$_1$) and i$_3$) (whence what we have stated for a group satisfying i$_0$), i$_1$), i$_2$), i$_3$) and ii) can certainly be applied to it). It is clear that i$_1$) is fulfilled. In order to verify i$_3$), we prove first

LEMMA 2. *Let K be a subring of our primary ring R which satisfies the condition* i$_2$)*; thus let $(a_1 = 1, a_l, \cdots, a_k)$ be an independent Z-basis of K such that $(a_{1r}, a_{2r}, \cdots, a_{kr})$ is an independent R_l-basis of R_lK_r. Then the endomorphisms of R/N induced by $a_{1r}, a_{2r}, \cdots, a_{kr}$ (i. e. $a_{1r}^0, a_{2r}^0, \cdots, a_{kr}^0$) are independent over the left-multiplication ring R_l^0 of R/N. In particular, $a_1 \bmod N, a_2 \bmod N, \cdots, a_k \bmod N$ are independent over the center of R/N, and the intersection of $(K+N)/N$ and the center of R/N is exactly $(Z+N)/N$. If we have* i$_3$)*, besides* i$_2$)*, then R/N is an $R_l^0K_r^0$-regular module.*

Proof. Suppose

$$\sum_{i=1}^{k} s_{il}^0 a_{ir}^0 = 0$$

for some elements s_i in R. We have $\sum_i s_ixa_i \,\varepsilon\, N$ for all x in R. Hence $l(N)\sum_i s_ixa_i = 0$ for all x in R, where $l(N)$ denotes the left annihilator of N in R. Since $a_{1r}, a_{2r}, \cdots, a_{kr}$ are independent over R_l, we have $l(N)s_i = 0$, for every $i = 1, 2, \cdots, k$. So each s_i is contained in the right annihilator $r(l(N))$ of $l(N)$. But $r(l(N)) = N$; this is in fact the only point where the primarity is used in our proof. Thus $s_i \,\varepsilon\, N$ $(i = 1, 2, \cdots, k)$, which proves the first assertion of our lemma. The second one follows easily from the first.

Now suppose that i$_3$) holds. Then the R_lK_r-right-module R_lK_r is isomorphic to R^k, the direct sum of k copies of R, and the R_lK_r-right-module $R_lK_r/N_lK_r = R_lK_r/(R_lK_r)N_l$ is isomorphic to $(R/RN_l)^k = (R/N)^k$. Here

$$\begin{aligned} R_lK_r/N_lK_r &\cong (R_l/N_l)a_{1r} \oplus (R_l/N_l)a_{2r} \oplus \cdots \oplus (R_l/N_l)a_{kr} \\ &\cong R_l^0a_{1r}^0 \oplus R_l^0a_{2r}^0 \oplus \cdots \oplus R_l^0a_{kr}^0 \end{aligned}$$

by the first half of our lemma. It follows that the $R_l^0K_r^0$-right-module $R_l^0K_r^0$ is isomorphic to $(R/N)^k$.

Next,

LEMMA 3. *Let a subring K of the primary ring R satisfy* i_2) *and suppose that the product* $((K+N)/N)$ (*center of* R/N) *is simple. Then R is R_lK_r-regular, i. e.* i_3) *is satisfied.*

Proof. Since R/N is centrally simple over its center, the product $R_l{}^0K_r{}^0$ (which is (homomorphic, whence) isomorphic to the direct (Kronecker) product $(R/N)\otimes((K+N)/N)$ (center of R/N)) over the center of R/N) is a simple ring. Thus its right-module R/N is $R_l{}^0K_r{}^0$-regular and $(R/N)^k$ is $R_l{}^0K_r{}^0$-isomorphic to $R_l{}^0K_r{}^0$, where k is the Z-rank of K (which is, according to Lemma 2, equal to the $R_l{}^0$-rank of $R_l{}^0K_r{}^0$). Now let u be an element of R^k such that $u \bmod N^k$ corresponds to the unit element of $R_l{}^0K_r{}^0$ by our $R_l{}^0K_r{}^0$-isomorphism of $(R/N)^k=R^k/N^k$ and $R_l{}^0K_r{}^0$. We consider the submodule $u(R_lK_r)$ of R^k. On mapping the unit element of R_lK_r onto u, we obtain an R_lK_r-homomorphism of R_lK_r onto $u(R_lK_r)$. If we pass to the residue-module $R^k/N^k=(R/N)^k$ the image $u(R_lK_r) \bmod N^k$ is all of R^k/N^k, since the mapping clearly induces our $R_l{}^0K_r{}^0$-isomorphism between $R_l{}^0K_r{}^0$ and R^k/N^k. However, as N^k is the intersection of all maximal R_l-submodules of R^k, it follows that $u(R_lK_r)=R^k$, too. Considering R_l-lengths, we conclude that our homomorphism mapping (which has thus been shown to be onto R^k) is an isomorphism between R_lK_r and R^k. This proves our lemma.

Remark. As our proof shows, our second assumption, in Lemma 3, may be replaced by the assumption that R/N be $R_l{}^0K_r{}^0$-regular.

We have thus seen that a regular automorphism group of R satisfies all the conditions in **1**.

Another remark we want to make is that if Φ is a regular automorphism group of our R then the automorphism group Φ^0 of the simple residue-ring R/N, induced by Φ, is a regular automorphism group of R/N, with the product $((K+N)/N)$ (center of R/N) as the associated simple subring of R/N generated by the elements of R/N effecting inner automorphisms of R/N contained in Φ^0. In fact, in order to see this we have only to observe that for $\sigma\neq 1$ the (minimal) R_l-two-sided modules R_l/N_l, $\rho(\sigma)R_l/\rho(\sigma)N_l$ are not isomorphic, or, which is the same, the $R_l{}^0$-two-sided modules $R_l{}^0$, $\rho(\sigma)^0R_l{}^0$ are not isomorphic, which is in turn equivalent to saying that $\rho(\sigma)^0$ is not an inner automorphism of the simple ring R/N, a consequence of ii).

It follows, either from I, Theorem 1 or from [5] that R/N possesses an independent right-basis of $(G:1)k$ terms over the invariant system, in R/N, of Φ^0. On the other hand, R has an invariant right-basis of $(G:1)k$ terms over the invariant system S, in R, of Φ. Here $(S+N)/N$ is clearly contained in the invariant system of Φ^0. It follows then that the invariant system of Φ^0 is exactly $(S+N)/N$. Hence $(S+N)/N$ is a simple ring. It follows further that an independent S-right-basis of R is independent over $(S+N)/N$ when considered modulo N. Thus we have

LEMMA 4. *Let Φ be a regular automorphism group of a primary ring R, possessing unit element and satisfying minimum condition, and let S be the invariant system of Φ in R. Then $(S+N)/N$, with N the radical of R, is the invariant system of the automorphism group Φ^0 of R/N induced by Φ. So $(S+N)/N$ is simple. (An independent S-right-basis of R exists and) an independent S-right-basis of R forms, when considered modulo N, an independent $(S+N)/N$-right-basis of R/N.*

Let S be the invariant system of a regular automorphism group Φ of our primary ring R. Consider a subring T of R which contains S;

$$R \supseteq T \supseteq S. \tag{5}$$

Suppose that $T+N/N$ is simple. As $(T+N)/N \supseteq (S+N)/N$ and $(S+N)/N$ is the invariant system of Φ^0 (as has been shown above) $(T+N)/N$ is a weakly normal simple subring of R/N, and hence R/N is $R_l{}^0T_r{}^0$ fully reducible; see [5], § 1. So, if we denote by $R^{(0)}$ the intersection of all maximal R_lT_r-submodules of R, then

$$N \supseteq R^{(0)}, \quad \text{whence} \quad l(N)R^{(0)} = 0. \tag{6}$$

Now we assert that none of $R^{\rho(\sigma)a_r}$ $(a \in K, a \notin N)$, σ running over G (cf. **1**, ii)), is contained in $R^{(0)}$. For, suppose $R^{\rho(\sigma)a_r} \subseteq R^{(0)}$. Then, with the left annihilator $l(N)$ of N, we have $l(N)R^{\rho(\sigma)a_r} = 0$, or $\rho(\sigma)a_r l(N)_l = 0$, $a_r l(N)_l = 0$, $l(N)Ra = 0$, whence $Ra \subseteq N$ and $a \in N$ contrary to the assumption. Naturally this applies to the case when a is any one of the basic elements $a_1, a_2, \cdots, a_k$ of K over Z.

Suppose further that R has an independent basis over our intermediate ring T. Then the T_r-endomorphism ring $V_{\mathfrak{A}}(T_r)$ of R has the form (4), $V_{\mathfrak{A}}(T_r) = \sum \rho(\sigma)R_lL_r{}^{(\sigma)}$, where each $R_lL_r{}^{(\sigma)}$ has an independent R_l-basis consisting of elements of $L_r{}^{(\sigma)}$ (Lemma 1). In particular, $L_r{}^{(1)}$ $(\rho(1) = 1)$ is $V_{\mathfrak{A}}(T) \cap R_r$ which is the R_lT_r-endomorphism ring of R;

$$L^{(1)} = V_R(T). \tag{7}$$

Put $L = L^{(1)}$ for the sake of simplicity. Then $L_r\rho(\sigma)R_lL_r{}^{(\sigma)}L_r \subseteq V_{\mathfrak{A}}(T_r)$ whence $\subseteq \rho(\sigma)R_lL_r{}^{(\sigma)}$,, and it follows that each $L^{(\sigma)}$ is L-two-sided allowable.

Now, decompose R into a direct sum of directly indecomposable R_lT_r-right-submodules $\mathfrak{m}_i$; $R = \mathfrak{m}_1 \oplus \mathfrak{m}_2 \oplus \cdots \oplus \mathfrak{m}_s$. Suppose that the $\mathfrak{m}_i$ are all mutually R_lT_r-isomorphic, or, which is the same, that $L = V_R(T)$ is primary; observe that $L_r = V_{\mathfrak{A}}(R_lT_r)$. On fixing a coherent system of isomorphisms among the $\mathfrak{m}_i$, we denote by $\{\epsilon_{ij}\}$ the corresponding system of matric units in the R_lT_r-endomorphism ring L_r of R. Let a be an element of K such that $a \notin N$. Then, as was shown above,

$$(8) \qquad R^{\rho(\sigma)a_r} \not\subseteq R^{(0)},$$

where $R^{(0)}$ is the intersection of all maximal R_lT_r-submodules of R. Hence the image by $\rho(\sigma)a_r$ of one, at least, of $\mathfrak{m}_1, \mathfrak{m}_2, \cdots, \mathfrak{m}_s$ is not contained in $R^{(0)}$, and we may assume, without loss in generality, that

$$(9) \qquad \mathfrak{m}_1^{\rho(\sigma)a_r} \not\subseteq R^{(0)}.$$

Then there exists an i such that the projection $\mathfrak{m}_1^{\rho(\sigma)a_r\epsilon_{ii}}$ of $\mathfrak{m}_1^{\rho(\sigma)a_r}$ on $\mathfrak{m}_i$ is not contained in $R^{(0)}$. We have, with such an i,

$$(10) \qquad \mathfrak{m}_1^{\rho(\sigma)a_r\epsilon_{ii}\epsilon_{i1}} \not\subseteq R^{(0)}.$$

Consider now a σ with $L^{(\sigma)} \neq 0$. Then $L^{(\sigma)} \not\subseteq N$; in fact $R_lL_r^{(\sigma)}$ has an independent basis over R_l consisting of elements of $L_r^{(\sigma)}$ (which is not void). So we may take an element a in $L^{(\sigma)}$ not contained in N, and apply our consideration to this a so as to obtain (10), with this a. There $\rho(\sigma)a_r\epsilon_{ii}\epsilon_{i1} = \rho(\sigma)a_r\epsilon_{i1}$ is an $(R_l, \rho(\sigma))$-semilinear and T_r-linear endomorphism of R; observe that $a_r, \epsilon_{i1} \varepsilon V_{\mathfrak{A}}(R_l)$ and $\rho(\sigma)a_r, \epsilon_{i1} \varepsilon V_{\mathfrak{A}}(T_r)$. It induces an endomorphism on $\mathfrak{m}_1$, which is not nilpotent. For, if it were, then the image would be contained in any maximal R_lT_r-submodule of R, whence in $R^{(0)}$, contrary to (10). By Fitting's lemma (in a generalized formulation as in Jacobson's book, Theory of Rings, Math. Surveys II, New York 1943, for example) $\rho(\sigma)a_r\epsilon_{i1}$ induces therefore an ($(R_l, \rho(\sigma))$-semilinear and T_r-linear) automorphism on $\mathfrak{m}_1$ (Let $R_l \cup T_r$ be the operator domain Ω in Jacobson's book).

As $\mathfrak{m}_1, \mathfrak{m}_2, \cdots, \mathfrak{m}_s$ are by assumption all $\{R_l, T_r\}$-isomorphic, it follows then that $R = \sum \mathfrak{m}_i$ has an $(R_l, \rho(\sigma))$-semilinear and T_r-linear automorphism, say α. Since $\alpha \varepsilon V_{\mathfrak{A}}(T_r)$ and αR_l $(\subseteq \mathfrak{A})$ is R_l-two-sided isomorphic with $\rho(\sigma)R_l$, we have

$$\alpha R_l \subseteq V_{\mathfrak{A}}(T_r) \cap \rho(\sigma)R_lK_r = \rho(\sigma)R_lL_r^{(\sigma)}.$$

So $\alpha R_lL_r \subseteq \rho(\sigma)R_lL_r^{(\sigma)}$. On the other hand, every element of $\alpha^{-1}\rho(\sigma)R_lL_r^{(\sigma)}$ is $\{R_l, T_r\}$-linear and it follows that $\alpha^{-1}\rho(\sigma)R_lL_r^{(\sigma)} \subseteq \rho(1)R_lL_r = R_lL_r$. Combining these two relations we have

$$(11) \qquad \rho(\sigma)R_lL_r^{(\sigma)} = \alpha R_lL_r.$$

Moreover, as $\alpha^{-1}\rho(\sigma)$ is contained in both R_lL_r and $V_{\mathfrak{A}}(R_l) = R_r$, we have $\alpha^{-1}\rho(\sigma) \varepsilon R_lL_r \cap R_r = L_r$ and $\alpha^{-1}\rho(\sigma) = b_r$ with an element b in L. Here b is regular (since $\alpha^{-1}\rho(\sigma)$ is a (module-) automorphism) and $b_l^{-1}\alpha^{-1}\rho(\sigma) = b_l^{-1}b_r$ is the inner ring-automorphism of R effected by b. Hence $b_l^{-1}\alpha^{-1}$ is a ring-automorphism of R. It is T_r-linear, i. e. it is a ring-automorphism of R leaving T elementwise fixed. Putting $\rho^*(\sigma) = \alpha b_l$, we have $\rho^*(\sigma) = \rho(\sigma)b_r^{-1}b_l \varepsilon \rho(\sigma)\Phi_0$ and $\rho^*(\sigma)$ is an element of Φ which is a representative of the coset $\rho(\sigma)\Phi_0$ of $\rho(\sigma)$ modulo Φ_0. Moreover $\rho(\sigma)R_lL_r^{(\sigma)} = \alpha R_lL_r = \rho^*(\sigma)b_l^{-1}R_lL_r = \rho^*(\sigma)R_lL_r$. Now, these hold for each σ with $L^{(\sigma)} \neq 0$. It is clear, from these considerations, that the totality of such σ form a subgroup, say, H, of G. So we have proved

LEMMA 5. *Let T be a subring of R containing the invariant system S of a regular automorphism-group Φ in R. Suppose that $(T + N)/N$ is simple, that R possesses an independent right-basis over T, and that the commuter $L = V_R(T)$ of T in R is primary. Then there exist a representative system $\{\rho^*(1) = 1, \rho^*(\sigma), \cdots, \rho^*(\tau)\}$ of Φ modulo Φ_0 and a subgroup H of G such that*

$$V_{\mathfrak{A}}(T_r) = \sum_{\sigma \varepsilon H} \rho^*(\sigma)R_lL_r. \tag{12}$$

L has an independent finite basis over Z such that the corresponding right multiplications on R constitute an independent R_l-basis of R_lL_r.

It is also clear that the subgroup Ψ of all elements of Φ leaving our T element-wise fixed is

$$\Psi = \bigcup_{\sigma \varepsilon H} \rho^*(\sigma)\Psi_0 \tag{13}$$

where Ψ_0 is the totality of inner automorphisms of R effected by regular elements of L, and

$$V_{\mathfrak{A}}(T_r) = R_l\Psi. \tag{14}$$

Further, $T_r = V_{\mathfrak{A}}(R_l\Psi)$ and T is the invariant system of Ψ. Thus

LEMMA 6. *Let T be as in Lemma 5, and let Ψ be the totality of automorphisms of R (or, of elements of Φ) leaving T element-wise fixed. Then T is the invariant system of Ψ in R.*

Now, Ψ clearly satisfies the conditions i$_0$), i$_2$) and ii), with K replaced by L (and Φ, Φ_0 by Ψ, Ψ_0, naturally). So Ψ is a regular automorphism-group if we assume that the product of $(L + N)/N$ and the center of R/N is simple. We observe also that this last condition implies that $((L + N)/N$ is simple and) L is primary. Thus

LEMMA 7. *Let T be a subring of R containing the invariant system S*

of a regular automorphism group Φ *in* R. *Suppose that* $(T+N)/N$ *is simple, that* R *has an independent right-basis over* T, *and that the product* $((L+N)/N)$ *(center of* R/N*) in* R/N *is simple, where* $L=V_R(T)$. *Then the totality* Ψ *of automorphisms of* R *(or, elements of* Φ*) leaving* T *element-wise fixed is a regular automorphism group of* R.

Next we start, conversely, with a regular subgroup Ψ of Φ, and let T be its invariant system. To Ψ and T we may naturally apply what we know about Φ and S. Thus Ψ exhausts the automorphisms of R leaving T element-wise fixed and (T is primary and) $(T+N)/N$ is simple. Combining this with Lemmas 5 and 6, we have our main

THEOREM 1. *Let* R *be a primary ring possessing a unit element and satisfying the minimum condition. Let* Φ *be a regular group of automorphisms of* R *and let* S *be the invariant system of* Φ *in* R. *Then (*S *is primary and)* $(S+N)/N$ *is simple, where* N *denotes the radical of* R. R *possesses an independent right-basis over* S, *the number of elements in the basis being equal to the reduced order of* Φ, *and* Φ *exhausts the automorphisms of* R *leaving* S *element-wise fixed. Further, the regular subgroups of* R *are in a 1-1 Galois correspondence with the intermediate rings* T *of* R *and* S *such that* R *possesses an independent right basis over* T, $(T+N)/N$ *is simple and, if* $L=V_R(T)$ *is the commuter of* T *in* R, *(*L *is primary and) the product* $((L+N)/N)$ *(center of* R/N*) is simple.*

For the sake of convenience, we repeat our definition of a regular automorphism group: a group Φ of automorphisms of R is called *regular* if the following conditions i), ii) are satisfied:

i) the totality Φ_0 of inner automorphisms of R contained in Φ is the totality of inner automorphisms of R effected by the regular elements in a certain subring K containing the center Z of R, generated by its regular elements, and possessing an independent finite basis $a_1, a_2, \cdots, a_k$ over Z (we may take $a_1=1$) such that $a_{1r}, a_{2r}, \cdots, a_{kr}$ are independent over R_l, and the product $((K+N)/N)$ (center of R/N) is simple, N being the radical of R,

ii) the factor group $G=\Phi/\Phi_0$ is finite, and, if $\{\rho(1), \rho(\sigma), \cdots, \rho(\tau)\}$ $(G=1, \sigma, \cdots, \tau)$ (we may take $\rho(1)=1$) is a representative system of Φ modulo Φ_0, the R_l-two-sided modules R_l, $\rho(\sigma)R_l$ $(\sigma\neq 1)$ have no isomorphic composition residue-modules.

The number $(G:1)(K:Z)$ is called the *reduced order* of the regular automorphism group Φ.

We note that our regular automorphism groups cross with the automorphism-groups considered in I in their generality, though the present groups are perhaps better qualified to be called Galois groups for primary rings.

3. Certain factor-groups of a regular automorphism group; extension of isomorphisms. Let Φ be an automorphism group of a ring R, with unit element and minimum condition, which satisfies the conditions i_0), i_1), i.), i_3), and ii). Consider a subgroup Ψ of Φ which satisfies similar conditions, L with $R_l\Psi \cap R_r = L_r$ replacing K and $\Psi_0 = \Phi_0 \cap \Psi$ replacing Φ_0. Let T be the invariant system in R of Ψ. R possesses over T an independent right basis of $(\Psi:\Psi_0)(L:Z)$ terms, and Ψ is the totality of automorphisms of R leaving T element-wise fixed. $V_{\mathfrak{A}}(T_r)$ has the form

$$R_l\Psi_0\rho^*(1) \oplus R_l\Psi_0\rho^*(\alpha) \oplus \cdots \oplus R_l\Psi_0\rho^*(\gamma),$$

where $\{\rho^*(1), \rho^*(\alpha), \cdots, \rho^*(\gamma)\}$ is a representative system, in Ψ, of the group $H = \{1, \alpha, \cdots, \gamma\} = \Psi/\Psi_0$ (which may be considered as a subgroup of $G = \Phi/\Phi_0$.

Assume now that Ψ is an invariant subgroup of Φ. Then T is mapped on itself by every element of Φ, and the factor-group Φ/Ψ is an automorphism group of T. We first prove

LEMMA 8. *Let ϕ be an element of Φ. If the T_l-two-sided modules T_l and ϕT_l have some isomorphic composition residue-modules, then $\phi \,\varepsilon\, \Phi_0\Psi$.*

Proof. For any T-double-module $\mathfrak{m}$ we can construct an R-T-double-module $R \otimes_T \mathfrak{m}$. Further, if ϕ is an automorphism of T, we may construct from $\mathfrak{m}$ a T-double-module $(\phi, \mathfrak{m})$ by crossing the operation of T on the left-hand side of $\mathfrak{m}$ with ϕ; thus $(\phi, \mathfrak{m})$ is identical with $\mathfrak{m}$ as T-right-module, while the left-operation of an element a of T on $(\phi, \mathfrak{m})$ coincides with that of a^ϕ on $\mathfrak{m}$. Now, if ϕ is induced by an automorphism of R, denoted again by ϕ, then we assert that the similarly defined R-T-double-module $(\phi, R \otimes_T \mathfrak{m})$ is (R-T-) isomorphic with $R \otimes_T (\phi, \mathfrak{m})$. In fact, this can easily be verified by taking an independent T-right-basis $(x_1, x_2, \cdots, x_h)$ of R and using the independent bases $(x_i^{\phi^{-1}})$ and (x_i) respectively in constructing $R \otimes_T (\phi, \mathfrak{m})$ and $(\phi, R \otimes_T \mathfrak{m})$; cf. [4], § 5.

We consider the case where $\mathfrak{m}$ is T itself and ϕ is an element of Φ. Naturally we have then $R \otimes_T \mathfrak{m} = R \otimes_T T = R$. Put $f = (\Psi:\Psi_0)(L:Z)$. Then R has an independent right-basis of f terms over T, and the direct sum R^f of f copies of R is $V_{\mathfrak{A}}(T_r)$-isomorphic with $V_{\mathfrak{A}}(T_r)$. The $V_{\mathfrak{A}}(T_r)$-endomorphism ring of R^f is the f-dimensional matric ring $(T_r)_f$ over T_r, while that of $V_{\mathfrak{A}}(T_r)$ is of course the left-multiplication ring of $V_{\mathfrak{A}}(T_r)$ which we shall denote by $\mathfrak{T}$. Hence there exists a (ring-) isomorphism λ between $(T_r)_f$ and $\mathfrak{T}$ which corresponds to a $V_{\mathfrak{A}}(T_r)$-isomorphism between R^f and $V_{\mathfrak{A}}(T_r)$. Thus R^f and $V_{\mathfrak{A}}(T_r)$ are $V_{\mathfrak{A}}(T_r)$-(linear) and $((T_r)_f - \mathfrak{T}, \lambda)$-semilinear isomorphic. They are in particular R_l-linear and $((T_r)_f - \mathfrak{T}, \lambda)$-semilinear isomorphic, for $R_l \subseteq V_{\mathfrak{A}}(T_r)$. We construct from R^f a second $\{R_l, (T_r)_f\}$-module $(R^f)^{(\phi)}$ by crossing the operator domain R_l with its automorphism ϕ. Similarly we construct $V_{\mathfrak{A}}(T_r)^{(\phi)}$. These two bodies are also R_l-(linear) and $((T_r)_f - \mathfrak{T}, \lambda)$-semilinear isomorphic.

Suppose now that the $\{R_l, T_r\}$-modules R and $R^{(\phi)}$ have some (properly linear) isomorphic composition residue-modules. Then the same is the case with the $\{R_l, (T_r)_f\}$-modules R^f and $R^{(\phi)f} = (R^f)^{(\phi)}$. In view of the above R_l-linear and $\{(T_r)_f - \mathfrak{T}, \lambda\}$-semilinear isomorphisms between R^f, $V_{\mathfrak{A}}(T_r)$ and between $(R^f)^{(\phi)}$, $V_{\mathfrak{A}}(T_r)^{(\phi)}$, we see that the $\{R_l, \mathfrak{T}\}$-modules $V_{\mathfrak{A}}(T_r)$ and $V_{\mathfrak{A}}(T_r)^{(\phi)}$ have some (properly linear) isomorphic composition residue-modules. This means however that the $V_{\mathfrak{A}}(T_r)$-R_l-double-modules $V_{\mathfrak{A}}(T_r)$, $V_{\mathfrak{A}}(T_r)\phi$ ($\subseteq \mathfrak{A}$) have some isomorphic composition residue-modules. Naturally the same is the case if we consider $V_{\mathfrak{A}}(T_r)$, $V_{\mathfrak{A}}(T_r)\phi$ merely as R_l-double-modules. But $V_{\mathfrak{A}}(T_r) = R_l\Psi = \sum_{\alpha \varepsilon H} R_l\Psi_0\rho^*(\alpha)$. Hence there are α, β in H such that $R_l\Psi_0\rho^*(\alpha)$, $R_l\Psi_0\rho^*(\beta)\phi$ have isomorphic composition residue-modules. It follows that $\Psi_0\rho^*(\alpha) \equiv \Psi_0\rho^*(\beta)\phi \bmod \Phi_0$ whence $\phi \varepsilon \Phi_0\Psi$.

Now, the assumption in our lemma, that the T_l-double-modules T_l, ϕT_l have some isomorphic composition residue-modules, is equivalent to assuming that the T-double-modules T and (ϕ, T) have some isomorphic composition residue-modules. This implies, because of our remark at the beginning of the present proof, that the R-T-double-modules R and (ϕ, R) have some isomorphic composition residue-modules. This is in turn equivalent to the existence of isomorphic composition residue-modules of the $\{R_l, T_r\}$-modules R and $R^{(\phi)}$, from which we have derived $\phi \varepsilon \Phi_0\Psi$. So our lemma is proved.

We have now immediately

THEOREM 2. *Let R be a ring, with unit element and minimum condition, and let Φ be an automorphism group of R satisfying the conditions* i_0), i_1), i_2), i_3) *and* ii), *with its subgroup Φ_0 as in* i_0). *(It is sufficient that R be primary and Φ a regular automorphism group of R). Let Ψ be an invariant subgroup of Φ which contains Φ_0 and let T be the invariant system of Ψ in R. Then Φ/Ψ is a regular automorphism-group of T containing no inner automorphism of T except the unity.*

(Automorphism groups of the type of this factor-group have been studied in [4], § 5 and the older paper in Jour. Math. Soc. Japan 1 (1949), pp. 203-216 cited there.)

Next we prove an extension theorem for isomorphisms of subrings (Theorem 3). To do so, we again consider an automorphism group Φ of a primary ring R, with unit element and minimum condition, which fulfills the conditions i_0), i_1), i_2), i_3) and ii) of 1. Let S be its invariant system. We have $V_{\mathfrak{A}}(S_r)=R_l\Phi=\sum_{\sigma\varepsilon G}\rho(\sigma)R_l\Phi_0$, $R_l\Phi_0=R_lK_r$. Let T be a subring of R which contains S and over which R has an independent right-basis. By Lemma 1, $V_{\mathfrak{A}}(T_r)$ has the form (3), i.e. $V_{\mathfrak{A}}(T_r)=\sum_{\sigma}\rho(\sigma)R_lL_r^{(\sigma)}$; each subring $L^{(\sigma)}$ of K contains Z and possesses an independent basis over Z such that the corresponding right-multiplications on R constitute an independent basis of $R_lL_r^{(\sigma)}$ over R_l.

Now, let α be a (ring-) isomorphic mapping of T into R which leaves S element-wise fixed. We assume that R has also an independent right-basis over T^α. Let $\{x_i\}$, $\{y_i\}$ be independent right-bases of R over T, T^α respectively; they consist of equal numbers of elements, since the T_r- and the T_r^α-ranks of R are equal as we readily see on considering S_r-ranks. The mapping

$$\sum x_it_i\rightarrow\sum y_it_i^\alpha\qquad(t_i\varepsilon T)\tag{15}$$

gives a $(T_r-T_r^\alpha,\alpha)$-semilinear automorphism μ of R. Let $\mathfrak{M}$ be the totality of $(T_r-T_r^\alpha,\alpha)$-semilinear endomorphisms of R; clearly $\mu\varepsilon\mathfrak{M}$. $\mathfrak{M}$ is an R_l-two-sided submodule of $V_{\mathfrak{A}}(S_r)$. So $\mathfrak{M}$ is a direct sum of submodules of $\rho(\sigma)R_lK_r$ $(\sigma\varepsilon G)$;

$$\mathfrak{M}=\sum_{\sigma}(\rho(\sigma)R_lK_r\cap\mathfrak{M}).$$

Further, $\mathfrak{M}\mu^{-1}\subseteq V_{\mathfrak{A}}(T_r)$ and $V_{\mathfrak{A}}(T_r)\mu\subseteq\mathfrak{M}$, whence

$$\mathfrak{M}=V_{\mathfrak{A}}(T_r)\mu=\sum\rho(\sigma)R_lL_r^{(\sigma)}\mu.\tag{16}$$

It follows that $\mathfrak{M}$ is R_l-left-regular. Its direct summands $\mathfrak{M}\cap\rho(\sigma)R_lK_r$ are then also R_l-left-regular. Lemma 0 implies that for each σ

$$\rho(\sigma)^{-1}\mathfrak{M}\cap R_lK_r=R_lM_r^{(\sigma)},\tag{17}$$

where $M^{(\sigma)}$ is, for each σ, a certain submodule of K (Apply Lemma 0 in its modified form with right-operations replaced by left-operations (but retaining R_l, K_r), or in its completely (right-left) symmetric form with R replaced by a ring inverse-isomorphic to R). Hence

$$\mathfrak{M}\cap\rho(\sigma)R_lK_r=\rho(\sigma)R_lM_r^{(\sigma)}.$$

Denote now by $\bar{R}^{(0)}$ the intersection of all the maximal $R_lT_r^\alpha$-submodules of R. As $\mathfrak{M}$ contains a (module-) automorphism μ, we have $R\mathfrak{M} = R$ and $R\mathfrak{M} \not\subseteq \bar{R}^{(0)}$. There exists, therefore, a $\tau \varepsilon G$ such that

$$(R\rho(\tau)R_lM_r^{(\sigma)} =)\ R(\mathfrak{M} \cap \rho(\tau)R_lK_r) \not\subseteq \bar{R}^{(0)}.$$

Fixing such a τ, we take an element m in $M^{(\sigma)}$ such

$$R\rho(\tau)m_r \not\subseteq \bar{R}^{(0)}. \tag{18}$$

Let

$$R = \mathfrak{m}_1 \oplus \mathfrak{m}_2 \oplus \cdots \oplus \mathfrak{m}_s \tag{19}$$

be a direct decomposition of R into directly indecomposable R_lT_r-submodules. There exists an index i such that $\mathfrak{m}_i\rho(\tau)m_r \not\subseteq \bar{R}^{(0)}$. We may, without loss of generality, assume

$$\mathfrak{m}_1\rho(\tau)m_r \not\subseteq \bar{R}^{(0)}. \tag{20}$$

Further, let

$$R = \bar{\mathfrak{m}}_1 \oplus \bar{\mathfrak{m}}_2 \oplus \cdots \oplus \bar{\mathfrak{m}}_{\bar{s}} \tag{21}$$

be a direct decomposition of R into directly indecomposable $R_lT_r^{\mathfrak{a}}$-submodules. There exists an index j such that the $\bar{\mathfrak{m}}_j$-projection, with respect to (21), of $\mathfrak{m}_1\rho(\tau)m$, is not contained in $\bar{R}^{(0)}$.

Next, let $\bar{\mathfrak{M}}$ be the totality of $(T_r^{\mathfrak{a}} — T_r, \alpha^{-1})$-semilinear endomorphisms of R, and let $R^{(0)}$ be the intersection of all the maximal R_lT_r-submodules of R. Clearly $\mu^{-1} \varepsilon \mathfrak{M}$, $\mathfrak{M} = \sum(\mathfrak{M} \cap \rho(\sigma)R_lK_r)$ and each $\bar{\mathfrak{M}} \cap \rho(\sigma)R_lK_r$ has the form $\rho(\sigma)R_l\bar{M}_r^{(\sigma)}$ with a subring $\bar{M}^{(\sigma)}$ of K. Let u be an element of $\bar{\mathfrak{m}}_j$ which is not contained in $\bar{R}^{(0)}$; observe that $\bar{\mathfrak{m}}_j \cap \bar{R}^{(0)}$ is the intersection of all the maximal $R_lT_r^{\mathfrak{a}}$-submodules of $\bar{\mathfrak{m}}_j$. Then we assert that

$$u\mu^{-1} \notin R^{(0)}. \tag{22}$$

To see this, we observe that μ, being an element of $\sum \rho(\sigma)R_lM_r^{(\sigma)}$, is a sum of elements of the form $\rho(\sigma)a_lm_r^{(\sigma)}$ $(a \varepsilon R, m^{(\sigma)} \varepsilon M^{(\sigma)})$. Since R is primary, we may choose the a's so that they are regular elements of R. Such a product $\rho(\sigma)a_lm_r^{(\sigma)}$, with a regular in R, is an R_l-semilinear endomorphism of R; the associated automorphism of R is the product of $\rho(\sigma)$ and the inner automorphism of R effected by a. As it is in $\mathfrak{M}$, it is also $(T_r — T_r^{\mathfrak{a}}, \alpha)$-semilinear. Hence it maps $R^{(0)}$ into $\bar{R}^{(0)}$. For, if we denote by U the counter-image of $\bar{R}^{(0)}$ by $\rho(\sigma)a_lm_r^{(\sigma)}$, then R/U is mapped R_l-semilinear and $(T_r — T_r^{\mathfrak{a}}, \alpha)$-semilinear isomorphically upon a certain $R_lT_r^{\mathfrak{a}}$-submodule of $R/\bar{R}^{(0)}$. As $R/R^{(0)}$ is $R_lT_r^{\mathfrak{a}}$-fully-reducible, it follows that R/U is R_lT_r-fully-reducible. Hence $U \supseteq R^{(\sigma)}$, or, which is the same, $\rho(\sigma)a_lm_r^{(\sigma)}$ maps $R^{(0)}$ into

$\bar{R}^{(0)}$. Therefore, μ itself maps $R^{(0)}$ into $\bar{R}^{(0)}$. Now, if we had $u\mu^{-1} \varepsilon R^{(0)}$, then $u = u\mu^{-1}\mu \varepsilon R^{(0)}\mu \subseteq \bar{R}$, which is a contradiction. So we have (22).

We have seen that there exists an index j such that the $\bar{\mathfrak{m}}_j$-projection of $\mathfrak{m}_1\rho(\tau)m_r$ is not contained in $\bar{R}^{(0)}$. Taking as our u (which has been assumed only to be an element of $\bar{\mathfrak{m}}_j$ not belonging to $\bar{R}^{(0)}$) an element in the $\bar{\mathfrak{m}}_j$-projection of $\mathfrak{m}_1\rho(\tau)m_r$ not contained in $\bar{R}^{(0)}$, we see that

$$(\bar{\mathfrak{m}}_j\text{-projection of } \mathfrak{m}_1\rho(\tau)m_r)\mu^{-1} \not\subseteq R^{(0)}.$$

As $\mu^{-1} \varepsilon \mathfrak{M} = \sum(\mathfrak{M} \cap \rho(\sigma)R_lK_r)$, there exist an element $\bar{\tau}$ in G and an element $\bar{m}$ in $\bar{M}^{(\tau)}$ such that

$$(\bar{\mathfrak{m}}_j\text{-projection of } \mathfrak{m}_1\rho(\tau)m_r)\rho(\tau)\bar{m}_r \not\subseteq R^{(0)}. \tag{23}$$

There exists then an index i such that the $\mathfrak{m}_i$-projection of the left-hand side of (23) is not contained in $R^{(0)}$.

Now, we assume that $\mathfrak{m}_1, \mathfrak{m}_2, \cdots, \mathfrak{m}_s$ are all R_lT_r-(linear) isomorphic and $\bar{\mathfrak{m}}_1, \bar{\mathfrak{m}}_2, \cdots, \bar{\mathfrak{m}}_{\bar{s}}$ are all $R_lT_r^{\alpha}$-isomorphic, or what is the same, both $V_R(T_r) = L^{(1)}$ and $V_R(T_r^{\alpha}) = \bar{L}^{(1)}$ are primary. Let $\{\epsilon_{hk}\}$, $\{\bar{\epsilon}_{pq}\}$ be the systems of matric unit R_lT_r- and $R_lT_r^{\alpha}$-endomorphisms of R with respect to (19) and (21), respectively. Naturally $\epsilon_{hk} \varepsilon L_r^{(1)}$ $(= (V_R(T))_r = V_{\mathfrak{A}}(R_lT_r))$ and $\epsilon_{pq} \varepsilon \bar{L}^{(1)}$. Our above result shows that, with suitable i, j, τ, $\bar{\tau}$, m $(\varepsilon M^{(\tau)})$ and $\bar{m}$ $(\varepsilon M^{(\bar{\tau})})$,

$$\mathfrak{m}_1\rho(\tau)m_r\bar{\epsilon}_{jj}\rho(\bar{\tau})\bar{m}_r\epsilon_{ii} \not\subseteq R^{(0)} \text{ whence } \not\subseteq \mathfrak{m}_i \cap R^{(0)}, \text{ and, therefore,}$$

$$m_1\rho(\tau)m_r\bar{\epsilon}_{jj}\rho(\bar{\tau})\bar{m}_r\epsilon_{i1} \not\subseteq \mathfrak{m}_1 \cap R^{(0)}. \tag{24}$$

Here $\rho(\tau)m_r\bar{\epsilon}_{jj}\rho(\bar{\tau})\bar{m}_r\epsilon_{i1}$ is $(R_l, \rho(\tau)\rho(\bar{\tau}))$-semilinear. It is, on the other hand, T_r-linear, for $\rho(\sigma)m_r$ $(\varepsilon \mathfrak{M})$ is $(T_r - T_r^{\alpha}, \alpha)$-semilinear and $\rho(\bar{\tau})\bar{m}_r$ $(\varepsilon \bar{\mathfrak{M}})$ is $(T_r^{\alpha} - T_r, \alpha^{-1})$-semilinear while $\bar{\epsilon}_{jj}$ and ϵ_{i1} are respectively T_r^{α}- and T_r linear. As in **2**, we apply Fitting's lemma, in its generalized formulation, to the directly indecomposable R_lT_r-module $\mathfrak{m}_1$. It asserts that $\rho(\tau)m_r\bar{\epsilon}_{jj}\rho(\bar{\tau})\bar{m}_r\epsilon_{i1}$ induces on $\mathfrak{m}_1$ either an $((R_l, \rho(\tau)\rho(\bar{\tau}))$-semilinear and T_r-linear) automorphism or a nilpotent endomorphism. If the latter were the case, then $\mathfrak{m}_1$ would be mapped, by $\rho(\tau)m_r\bar{\epsilon}_{jj}\rho(\bar{\tau})\bar{m}_r\epsilon_{i1}$, into each maximal R_lT_r-submodule of $\mathfrak{m}_1$, whence into $\mathfrak{m}_1 \cap R^{(0)}$, contrary to (24). Hence $\rho(\tau)m_r\bar{\epsilon}_{jj}\rho(\bar{\tau})\bar{m}_r\epsilon_{i1}$ induces an automorphism on $\mathfrak{m}_1$.

It follows that $\rho(\tau)m_r\bar{\epsilon}_{jj}$ maps $\mathfrak{m}_1$ isomorphically into $\bar{\mathfrak{m}}_j$. As $\rho(\tau)m_r\bar{\epsilon}_{jj}$ is $(T_r - T_r^{\alpha}, \alpha)$-semilinear, whence S_r-linear, we see that the S_r-rank of $\mathfrak{m}_1$ is at most equal to the S_r-rank of $\bar{\mathfrak{m}}_j$, which is naturally equal to that of $\bar{\mathfrak{m}}_1$. Interchanging the roles of T, T^{α} we see similarly that the S_r-rank of $\bar{\mathfrak{m}}_1$ is at most equal to that of $\mathfrak{m}_1$. Hence they are equal and $\rho(\tau)m_r\bar{\epsilon}_{jj}$ induces an

$(R_l, \rho(\tau))$-semilinear and $(T_r - T_r^{\alpha}, \alpha)$-semilinear isomorphism of $\mathfrak{m}_1$ onto $\bar{\mathfrak{m}}_j$. As $\bar{\mathfrak{m}}_j$ is $R_l T_r^{\alpha}$-isomorphic to $\bar{\mathfrak{m}}_1$, we have thus established the existence of an $(R_l, \rho(\tau))$-semilinear and $(T_r - T_r^{\alpha}, \alpha)$-semilinear isomorphism between $\mathfrak{m}_1$ and $\bar{\mathfrak{m}}_1$ (or between any $\mathfrak{m}_k$ and any $\bar{\mathfrak{m}}_p$). We now find readily that the numbers s, $\bar{s}$ in (19), (21) are equal. For each $k = 1, 2, \cdots, s$ there exists an $(R_l, \rho(\tau))$-semilinear and $(T_r - T_r^{\alpha}, \alpha)$-semilinear isomorphism between $\mathfrak{m}_k$, $\bar{\mathfrak{m}}_k$ and there exists therefore a (module-) automorphism of R of the same kind, say ϕ. As $\phi^{-1} R_l \phi = R_l$ and $R_r = V_{\mathfrak{A}}(R_l)$ we have $\phi^{-1} R_r \phi = R_r$ and the inner automorphism of $\mathfrak{A}$ effected by its regular element ϕ induces a (ring)-automorphism of R_r. It induces on T_r the isomorphism α with T^{α}. This shows that the isomorphism α between T and T^{α} is extended to a (ring-) automorphism of R. The extended automorphism belongs to Φ, since it leaves S element-wise fixed. So we have

THEOREM 3. *Let Φ be an automorphism group of a primary ring R, with unit element and minimum condition, which satisfies the conditions* $i_0)$, $i_1)$, $i_2)$, $i_3)$ *and* ii) (*of* 1), *and let S be its invariant system* (It is sufficient that Φ be a regular automorphism group of R). *Let T be a subring of R containing S and having a primary commuter $V_R(T)$ in R, and let α be an isomorphic mapping of T into R leaving S element-wise fixed such that the commuter $V_R(T^{\alpha})$ of the image T^{α} is also primary. Then α can be extended to an element of Φ.*

NAGOYA UNIVERSITY.

REFERENCES,

[1] H. Cartan, "Théorie de Galois pour les corps non-commutatifs," *Annales Scientifiques de l'École Normale Superiere*, vol. 65 (1948), pp. 60-77.

[2] N. Jacobson, "A note on division rings," *American Journal of Mathematics*, vol. 69 (1947), pp. 27-36.

[3] T. Nakayama, "Commuter systems in a ring with radical," *Duke Mathematical Journal*, vol. 16 (1947), pp. 331-337.

[4] ———, "Automorphisms of simple, complete primitive, and directly indecomposable rings," *Annals of Mathematics*, vol. 55 (1952), pp. 538-551.

[5] ———, "Galois theory of simple rings," *Transactions of the American Mathematical Society*, vol. 73 (195–), pp. 276-292.

PERIODIC MAPPINGS ON BANACH ALGEBRAS.*

By Bertram Yood.

---

1. Introduction. Let B be a real Banach algebra. In investigations of such algebras, those with an involution (adjoint operation) have received special attention. By an involution is meant an anti-automorphism of period two. Here we study automorphisms and anti-automorphisms T of B periodic with finite period (the period of T is the smallest positive integer n such that $T^n = I$, where I is the identity).

The norms $\| x \|$ and $\| x \|_1 = \| T(x) \|$ are two Banach algebra norms for B which are equivalent if and only if T is continuous ([1], p. 41). The equivalence of two Banach algebra norms for B has been discussed by Rickart [10] where it is shown that a necessary and sufficient condition for the equivalence is that $S = (0)$ where S is the separating ideal of the two norms, the set of elements x where for some sequence $\{x_n\}$ in B, $\| x_n - x \| \to 0$ and $\| x_n \|_1 \to 0$. As shown in [10], S is a closed two-sided ideal of B consisting entirely of topological divisors of zero. A further property of S is (Theorem 2.4) that either the two norms are equivalent or S is a radical algebra in the sense of the Brown-McCoy radical [2]. Our approach to the question of the continuity of T is made through the study of the corresponding ideal S.

It has been conjectured [10] that any two Banach algebra norms for a semi-simple algebra B are equivalent. This would imply the continuity of any periodic T defined on B. In this direction we show, in Theorem 4.6, that if every automorphism and anti-automorphism T of period two defined on each closed semi-simple sub-algebra B_1 of B is continuous then any T of period 2^n defined on B is continuous. Furthermore if B is semi-simple, the period n of $T = 3$ or 2^m and if the set $\{x \varepsilon B \mid (I + T + \cdots + T^{n-1})(x) = 0\}$ is closed, then T is continuous.

2. Definitions and preliminaries. Let T be a linear mapping defined on an algebra B over the real field whose range is B and which has period n. Suppose that, for each x, y in B either $T(xy) = T(x)T(y)$ or $T(xy) = T(y)T(x)$. By a result of Hua [4] one of these two equations holds throughout B and T is either an automorphism or an anti-automorphism. Let

$$P_n = (I + T + \cdots + T^{n-1})/n$$

---

* Received April 26, 1954.

where I is the identity. Let

$$H = \{x \varepsilon B \mid T(x) = x\} \quad \text{and} \quad K = \{x \varepsilon B \mid P_n(x) = 0\}.$$

If $P_n(x) = x$ then $x = TP_n(x) = T(x)$, $x \varepsilon H$ and conversely $x \varepsilon H$ implies that $P_n(x) = x$. It is then clear that P_n is a projection of B on H and that H and K are complementary linear manifolds, $B = H \oplus K$.

2.1. Lemma. *$T(H) = H$. $T(K) = K$. $T^i(x) - x \varepsilon K$ for all $x \varepsilon B$ and all positive integers i.*

The first statement is clear. If $x \varepsilon K$ then $TP_n(x) = P_nT(x) = 0$ and $T(K) \subset K$. By iteration we obtain $T^{-1}(K) = T^{n-1}(K) \subset K$. Thus $T(K) = K$. If $y = T^i(x) - x$ then a simple computation shows that $P_n(y) = 0$.

Now for each integer i, $i = 1, \cdots, n-1$, T^i is also periodic. We define H_i and K_i as the sets H and K above corresponding to the mappings T^i.

2.2. Lemma. *Let T have period n. Then for $i = 1, 2, \cdots, n-1$ the following hold.*

(a) $H_1 \subset H_i$, $K_i \subset K_1$.

(b) $H_i = H_1 \oplus H_i \cap K_1$, $K_1 = K_i \oplus H_i \cap K_1$.

(c) *$H_1 = H_i$ if and only if $K_1 = K_i$.*

(d) *If i divides n then $H_i \cap K_1$ is the null-space of $I + T + \cdots + T^{i-1}$.*

(e) *If i divides n then $H_1 = H_i$ if and only if $I + T + \cdots + T^{i-1}$ is one-to-one.*

(a). Let $y \varepsilon K_i$, α be the greatest common divisor of i and n and set $c = n/\alpha$. By the definition of K_i we obtain

$$y + T^{\alpha}(y) + \cdots + T^{(c-1)\alpha}(y) = 0. \tag{1}$$

If we operate on (1) by $T, T^2, \cdots, T^{\alpha-1}$ and add we obtain $P_n(y) = 0$ or $y \varepsilon K_1$.

(b). It is clear that $H_i \supset H_1 \oplus H_i \cap K_1$. Let $x \varepsilon H_i$, $x = u + v$, $u \varepsilon H_1$, $v \varepsilon K_1$. Then $x = T^i(x) = u + T^i(v) = u + v$. Thus $v \varepsilon H_i \cap K_1$. For the second formula of (b) note that

$$B = H_i \oplus K_i = H_1 \oplus [K_i \oplus H_i \cap K_1] = H_1 \oplus K_1$$

where, by (a), $K_i \oplus H_i \cap K_1 \subset K_1$. Thus $K_1 = K_i \oplus H_i \cap K_1$.

(c). $H_i = H_1$ (or $K_i = K_1$) if and only if $H_i \cap K_1 = 0$.

(d). Let $n = ci$. Set $V = I + T + \cdots + T^{i-1}$ and $W = I + T^i + T^{2i} + \cdots + T^{(c-1)i}$. When $WV = VW = I + T + \cdots + T^{n-1}$. If $x \varepsilon H_i \cap K_1$ then $W(x) = cx$. $VW(x) = 0$ so that $V(x) = 0$. Conversely if $V(x) = 0$ then $WV(x) = 0$ so that $x \varepsilon K_1$. Also $(TV - V)(x) = 0$ where $TV - V = T^i - I$. Thus $x \varepsilon H_i$.

(e). This follows from (b) and (d).

2.3. Lemma. *Let* $v \varepsilon K_1$, $y \varepsilon H_1$. *Then* $vy + yv \varepsilon K_1$.

Note that $T^i(vy + yv) = T^i(v)y + yT^i(v)$. Thus $P_n(vy + yv) = 0$.

Let B be a real Banach algebra which is a Banach algebra under two norms $\| x \|$ and $\| x \|_1$. An element s in B is said ([10], p. 619) to *separate* the norms if there exists a sequence x_n in B such that $\| s - x_n \| \to 0$ and $\| x_n \|_1 \to 0$. The set S of all elements which separate the two norms has been shown by Rickart ([10], p. 621) to be a two-sided ideal in B, closed with respect to both norms and each of whose elements is a topological divisor of zero with respect to both norms. S will be called the *separating ideal* for these two norms. The same separating ideal is obtained if the roles of $\| x \|$ and $\| x \|_1$ are interchanged. The two norms are equivalent if and only if S is the zero element.

Consider next an automorphism or anti-automorphism U of period n. Define $\| x \|_1 = \| U(x) \|$. Following ideas of Rickart ([9], p. 1068) we show that $\| x \|_1$ is a Banach algebra norm for B.

It is clear that the norm makes B_1 into a normed linear space. Moreover $\| xy \|_1 = \| U(xy) \| \leqq \| U(x) \| \, \| U(y) \| = \| x \|_1 \| y \|_1$. Also if $\| x_m - x_n \|_1 \to 0$ as $m, n \to \infty$ then there exists $y \varepsilon B$ such that $\| y - U(x_n) \| \to 0$. Then $\| U^{-1}(y) - x_n \|_1 \to 0$ so that B is Banach algebra under the norm $\| x \|_1$. The separating ideal S for the two norms $\| x \|$ and $\| x \|_1$ is called the *separating ideal for* U.

If S is the separating ideal corresponding to two Banach algebra norms $\| x \|$ and $\| x \|_1$ for B then $\| x \|$ and $\| x \|_1$ are two Banach algebra norms on S and thus there is a separating ideal $S_1 \subset S$ corresponding to the two norms.

2.4. Theorem. *Let* $\| x \|$ *and* $\| x \|_1$ *be two Banach algebra norms for* B *with separating ideal* S *and with separating ideal* S_1 *with respect to* S. *Then* S *and* S_1 *have the following properties.*

(1) $x_1 x_2 \varepsilon S_1$ *if* $x_i \varepsilon S$, $i = 1, 2$. S_1 *is a two-sided ideal of* B.

(2) S_1 *is contained in no proper regular (one-sided or two-sided) ideal of* S.

(3) *Either the two norms are equivalent on B or S is a radical algebra in the Brown-McCoy sense.*

(4) *If the two norms are equivalent on S, then S is a zero algebra.*

(1). *Let $y_n \varepsilon B$, $\| y_n - x_1 \| \to 0$, $\| y_n \|_1 \to 0$.* Then $\| y_n x_2 - x_1 x_2 \| \to 0$ and $\| y_n x_2 \|_1 \to 0$ where $y_n x_2$ and $x_1 x_2 \varepsilon S$. Thus $x_1 x_2 \varepsilon S_1$. The second statement is shown in a similar way.

(2). Let J be, say, a proper regular right ideal in S with e as a left identity of S modulo J. If $S_1 \subset J$ then, since $ex - x \varepsilon J$ for all $x \varepsilon S$ and $ex \varepsilon S_1$ by (1), we have $S \subset J$ which is impossible.

(3). Suppose that the two norms are not equivalent. Then $S \neq (0)$. From the results of Brown and McCoy ([2], p. 51) we see that S is a radical algebra in the sense of their radical if and only if S contains no maximal regular two-sided ideal. However if M is such an ideal in S, results of Rickart ([10], p. 621) show that $S_1 \subset M$. By (2) the latter is impossible.

(4). If the two norms are equivalent on S then $S_1 = (0)$ and (4) then follows from (1).

As applied to a commutative Banach algebra B, Theorem 2.4 shows that S is then a radical algebra contained in the radical of B. This follows since, in the commutative case, the Brown-McCoy radical and the Jacobson radical ([3], [5]) agree.

2.5 LEMMA. *Let A be an algebra over the complex field with (Jacobson) radical R. Let R_1 be its radical when A is considered as an algebra over the real field. Then $R = R_1$.*

It is clear from the definitions ([3], p. 476) that $R \subset R_1$. Let $x \varepsilon R_1$, c real, $\neq 0$. Then cx^2 has an adverse in A. This shows that cix has an adverse in A, for otherwise $-i/c$ is in the spectrum of x ([3], p. 458) and $-c^{-2}$ is in the spectrum of x^2 which is impossible. Let $\alpha = a + bi$, a, b real, be any complex number. Let r be the adverse of bix. Using the notation $t_1 \circ t_2 = t_1 + t_2 - t_1 t_2$, we have, by [3], p. 477, for any $s \varepsilon A$,

$$[bix + (ax + yx)] \circ (r \circ s) = [ax + yx - (ax + yx)r] \circ s.$$

Now $ax + yx - (ax + yx)r \varepsilon R_1$ and thus has a right adverse s. This shows that $\alpha x + yx$ has a right adverse. Similar arguments ([3], p. 477) show that $\alpha x + yx$ has a left adverse. Thus $x \varepsilon R$.

This shows that the results of Section 4 concerning semi-simple real Banach algebras hold for semi-simple complex Banach algebras.

3. Examples. We indicate one way in which our mappings of period $n > 2$ can behave differently from those of period 2. First consider the following easy result where T is an automorphism or anti-automorphism of finite period on a real Banach algebra B.

3. 1. Lemma. *T is continuous on B if and only if H_1 and K_1 are closed and T is continuous on K_1.*

Define for $x = u + v$, $u \varepsilon H_1$, $v \varepsilon K_1$ a new norm $\| x \|_1 = \| u \| + \| v \|$. If H_1 and K_1 are closed this is readily seen to be a complete Banach space norm for B. Since $\| x \|_1 \geqq \| x \|$, the two norms give the same topology for B by [1], p. 41. Next assume that T is continuous on K_1. There is a constant $c > 0$ such that $\| T(v) \| \leqq c \| v \|$, $v \varepsilon K_1$. Then, for $x = u + v$, as above,

$$\| T(x) \|_1 = \| u \| + \| T(v) \| \leqq (1 + c) \| x \|_1$$

which gives the continuity of T.

Note that the continuity of T on K_1 is automatic if the period of T is two, for then $T(x) = -x$, $x \varepsilon K_1$. However if the period exceeds two it is possible, as the following example shows, to have H_1, K_1 closed while T is not continuous on K_1.

Let B be the complex L_p space on the unit interval where $p > 1$ and $p \neq 2$. We consider B as a zero algebra. As is well-known [8] there exists a closed complex linear manifold M in B on which there exists no continuous complex-linear projection of B. Let N be a complex linear manifold which is a complementary manifold to M. For $x = u + v$, $u \varepsilon M$, $v \varepsilon N$ define T by the rule $T(x) = \omega u + \omega^2 v$ where ω and ω^2 are the non-real cube roots of unity. It is easy to check that T is linear on B and that $I + T + T^2 = 0$. Then $T^3 = I$ and T is periodic. Here $K_1 = B$ and $H_1 = (0)$.

Now N is not closed for otherwise, [8], Lemma 1. 1. 1, there exists a continuous complex-linear projection of B on M. Let $x_0 \varepsilon \bar{N}$, $x_0 \notin N$, $x_0 = u_0 + v_0$, $u_0 \varepsilon M$, $v_0 \varepsilon N$. Then $u_0 \neq 0$, $u_0 \varepsilon M \cap \bar{N}$. Let $v_n \varepsilon N$, $v_n \rightarrow u_0$. We have $T(v_n) = \omega^2 v_n \rightarrow \omega^2 u_0$ while $T(u_0) = \omega u_0$. Thus T is discontinuous.

Next we show that H_1 can be closed with K_1 not closed. In this example we do not use a zero algebra. Let B be the L_p space considered above with, however, a different rule for multiplication. Let f be a linear functional on B, $\| f \| = 1$ such that f vanishes on M. Select $W \varepsilon M$, $\| w \| = 1$. Define the product in B by the rule $x_1 x_2 = f(x_1) f(x_2) w$. It is easy to verify that this definition makes B into a commutative Banach algebra. Consider the mapping $x \rightarrow x^*$ defined by the rule that if $x = u + v$, $u \varepsilon M$, $v \varepsilon N$ then $x^* = u - v$. This mapping is an automorphism of B of period two. We have $H = M$,

214915

$K = N$ and thus H is closed whereas K is not. Using these ideas an example can also be given where K is closed and H is not. As shown in Section 4 this type of behavior cannot take place if B is semi-simple.

4. The separating ideal and continuity for T. Henceforth T will represent a periodic automorphism or anti-automorphism with finite period on a real Banach algebra B with separating ideal S.

4.1. LEMMA. (1) $S \subset \bar{K}_1$. (2) $T(S) = S$. (3) *If* $u + v \varepsilon S$ *where* $u \varepsilon H_1$, $v \varepsilon K_1$ *then* $u \varepsilon S$ *and* $v \varepsilon S$. (4) $S = S \cap H_1 \oplus S \cap K_1$.

(1). Let $y \varepsilon S$, $x_k \to y$, $T(x_k) \to 0$. Set $x_k = u_k + v_k$ where $u_k \varepsilon H_1$ and $v_k \varepsilon K_1$. Then $u_k + T(v_k) \to 0$. This shows that $v_k - T(v_k) \to y$. But $v_k - T(v_k) \varepsilon K_1$ by Lemma 2. 1.

(2). Since $T^{n-1} = T^{-1}$, where n is the period of T, it is enough to show that $T(S) \subset S$. Let $y = u + v \varepsilon S$ where $u \varepsilon H_1$, $v \varepsilon K_1$. Adopt the notation of the proof of (1). Since $u_k + v_k \to u + v$ and $u_k + T(v_k) \to 0$, we have $(u - u_k) + (v - v_k) \to 0$ and $T[(u - u_k) + (v - v_k)] \to u + T(v) = T(y)$. Thus $T(y) \varepsilon S$.

(3). By (2), $P_n(S) \subset S$. Thus, in the above notation, $P_n(y) = u \varepsilon S$. Then also $v = y - u \varepsilon S$.

(4). This follows from (3).

Below the notation Q' is used for the derived set of Q.

4. 2. LEMMA. *Let T have period n on B. Then*

(a) $H_1' \cap K_1 \subset S$.

(b) *n even, T^2 continuous on $\bar{K}_1$ imply that* $K_1' \cap H_1 \subset S$.

(c) *$n = 2$ implies that* $S = H_1' \cap K_1 \oplus K_1' \cap H_1$.

(a). Let $u_k \varepsilon H_1$, $u_k \to v$, $v \varepsilon K_1$. Then $u_k - v \to 0$ and $T(u_k - v) \to v - T(v)$. Thus $v - T(v) \varepsilon S$. Then, by Lemma 4. 1, $T^i(v) - T^{i+1}(v) \varepsilon S$, $i = 0, 1, 2, \cdots$. Also

$$\sum_{m=0}^{n-1} \sum_{i=0}^{m} [T^i(v) - T^{i+1}(v)] = \sum_{m=0}^{n-1} [v - T^{m+1}(v)] = nv.$$

Therefore $v \varepsilon S$.

(b). Let $v_k \varepsilon K_1$, $v_k \to u$, $u \varepsilon H_1$. Set $W = I + T^2 + T^4 + \cdots + T^{n-2}$. Then $TW(v_k) = - W(v_k)$. Also $W(v_k) \to (n/2)u$. Then $(n/2)u + W(v_k) \to nu$, $T[(n/2)u + W(v_k)] \to 0$. Thus $u \varepsilon S$.

(c). The conclusions (a) and (b) show that here

$$H_1' \cap K_1 \oplus K_1' \cap H_1 \subset S.$$

Let $y \varepsilon S$. Then there exist sequences $\{u_k\}$ and $\{v_k\}$ in H_1 and K_1 respectively such that $u_k + v_k \to y$, $u_k - v_k \to 0$. Thus $y/2 = \lim u_k = \lim v_k$. Via Lemma 4.1, $S = S \cap H_1 \oplus S \cap K_1 \subset K_1' \cap H_1 \oplus H_1' \cap K_1$.

4.3. THEOREM.

(a) *If $S \subset H_1$ then H_1 is closed and S is a zero algebra.*

(b) *If H_1 is closed and T is continuous on K_1 then $S \subset H_1$.*

(a). Let $v \varepsilon H_1' \cap K_1$. By Lemma 4.2, $v \varepsilon S \subset H_1$. Thus $v \varepsilon H_1 \cap K_1 = (0)$. *Apply Theorem* 2.4.

(b). Let n be the period of T. In view of Lemma 4.1 it is enough to show that $S \cap K_1 = (0)$. Let $y \varepsilon S \cap K_1$, $u_k \varepsilon H_1$, $v_k \varepsilon K_1$ where $u_k + v_k \to y$ and $u_k + T(v_k) \to 0$. Then $v_k - T(v_k) \to y$. By Lemma 2.1, $T^i(v_k) - T^{i+1}(v_k) \varepsilon K_1$, $T^i(y) \varepsilon K_1$. By the continuity of T on K_1 we obtain

$$(1) \qquad T^i(v_k) - T^{i+1}(v_k) \to T^i(y), \qquad i = 0, 1, 2, \cdots.$$

But, as in the proof of Lemma 4.2,

$$(2) \qquad \sum_{m=0}^{n-1} \sum_{i=0}^{m} [T^i(v_k) - T^{i+1}(v_k)] = nv_k.$$

From (1) and (2) it follows that

$$(3) \qquad nv_k \to \sum_{m=0}^{n-1} \sum_{i=0}^{m} T^i(y) = ny + \sum_{m=1}^{n-1} (n-m) T^m(y)$$

$= w$, say. Then $nu_k \to ny - w \varepsilon K_1$. Since H_1 is closed, $w = ny$. This gives

$$(4) \qquad \sum_{m=1}^{n-1} (n-m) T^m(y) = 0.$$

Operating on this by T^{-1} we obtain

$$(5) \qquad \sum_{m=0}^{n-2} (n-m-1) T^m(y) = 0.$$

Subtracting (4) from (5) we have

$$(6) \qquad 0 = (n-1)y - \sum_{m=1}^{n-1} T^m(y) = ny.$$

This completes the proof.

4.4. LEMMA. *Let T have even period n with separating ideal S. Let S_1 be the separating ideal for T on S. Then*

(a) *$S \subset K_1$ and T^2 continuous on $S(\bar{K}_1)$ imply that $S_1 \subset K_2$ (K_1 closed).*

(b) *If, in addition, $n = 2$ (4) then $S(S_1)$ is a zero algebra.*

(a). Let n be the period of T. Suppose $y \varepsilon S_1$. Then there exists a sequence $\{v_k\} \subset S$, $v_k \rightarrow y$, $T(v_k) \rightarrow 0$. By the continuity of T^2 on S we obtain $T^i(v_k) \rightarrow T^i(y)$ or zero according as i is even or odd. Then, since $v_k \varepsilon K_1$, $(I + T^2 + \cdots + T^{n-2})(y) = 0$ or $y \varepsilon K_2$.

Now suppose that T^2 is continuous on $\bar{K}_1$. By Lemma 4.2, $K_1' \cap H_1 \subset S \subset K_1$. Thus $K_1' \cap H_1 = (0)$.

(b). If $n = 2$, then since here $K_2 = (0)$, S is a zero algebra by Theorem 2.4. If $n = 4$, then T^2 is an automorphism of period two and, for $x, y \varepsilon S_1$, $-xy = T^2(xy) = T^2(x)T^2(y) = xy$.

The following algebraic result may have some independent interest.

4.5. LEMMA. *Let B be a semi-simple algebra where T is an automorphism of period 2^n, $n = 0, 1, 2, \cdots$. Then H_1 is a semi-simple algebra.*

Since T is an automorphism, H_1 is a sub-algebra of B. The proof is by induction. If $n = 0$ then $H_1 = B$ and the conclusion trivial. Suppose that the conclusion holds for n. Let T have period 2^{n+1}. Then T^2 has period 2^n. By induction hypothesis, H_2 is semi-simple. If $H_2 = H_1$ we are through. Otherwise consider the decomposition $H_2 = H_1 \oplus H_2 \cap K_1$ provided by Lemma 2.2 where $H_2 \cap K_1 = \{x \varepsilon B \mid T(x) = -x\}$. T has period two as an automorphism defined on H_2. Let x be in the radical R of H_1. Then $\alpha x + yx$ has an adverse (reverse) in H_1 for all scalars α and all $y \varepsilon H_1$ (see [3], Chapter 22). Consider $z \varepsilon H_2 \cap K_1$. It is easy to see that $(zx)^2$ and $zxz \varepsilon H_1$. Thus $(zx)^2 \varepsilon R$. Using the notation $t_1 \circ t_2 = t_1 + t_2 - t_1 t_2$ we have, for every $s \varepsilon B$ the identities

$$(zx) \circ [(-zx) \circ s] = (zx)^2 \circ s, \qquad [s \circ (-zx)] \circ (zx) = s \circ (zx)^2.$$

If s is taken as the adverse of $(zx)^2$ in H_1 it follows from [3], p. 456, that $(-zx) \circ s$ is the adverse of zx. Moreover since $-zx \varepsilon H_2 \cap K_1$, this adverse lies in H_2.

Let $w = u + v$, $u \varepsilon H_1$, $v \varepsilon H_2 \cap K_1$ be any element of H_2 and let α be any scalar. $\alpha x + ux$ has an adverse p in H_1 since $x \varepsilon R$. Since $x - xp \varepsilon R$ then $v(x - xp)$ has an adverse q in H_2 by the above. But then (see [3], p. 477) we have, where $\alpha x + wx = (\alpha x + ux) + vx$, the relation

$$(\alpha x + wx) \circ (p \circ q) = (vx - vxp) \circ q = 0.$$

Also $v - pv \varepsilon H_2 \cap K_1$ so that $(v - pv)x$ has an adverse q_1 in H_2 and $(q_1 \circ p) \circ (\alpha x + wx) = 0$. Thus x lies in the radical of H_2. Hence H_1 is semi-simple and the induction is complete.

As applied to an anti-automorphism T of period 2^n on B this result gives the semi-simplicity of H_2.

4. 6. THEOREM. *Let B be a semi-simple Banach algebra. Then any T with period 2^n defined on B such that its separating ideal $S \subset K_1$ (as when K_1 is closed) is continuous.*

If $n = 1$ then $T^2 = I$ and S is a zero algebra by Lemma 4. 4 so the result holds for $n = 1$ by the semi-simplicity of B. For $n > 1$ we argue as follows. We show, by induction, that $S \subset K_{2^i}$ for $i = 0, 1, \cdots, n$. This is true for $i = 0$ by assumption. Suppose it is true for i, $0 \leqq i < n$. Now $H_{2^{i+1}}$ is semi-simple (Lemma 4. 5). Also $H_{2^{i+1}} \cap K_{2^i} = \{x \varepsilon B \,|\, T^{2^i}(x) = -x\}$ by Lemma 2. 2. Let $J = S \cap H_{2^{i+1}}$. J is a two-sided ideal in $H_{2^{i+1}}$. Let $u_k \varepsilon J$, $k = 1, 2$. Then $u_1 u_2 \varepsilon J$. If $i = 0$, $T^{2^i} = T$ may be either an automorphism or anti-automorphism. If $i > 0$, T^{2^i} is an automorphism. We have then either $-u_1u_2 = T^{2^i}(u_1u_2) = u_1u_2$ or $-u_1u_2 = u_2u_1$. In any case $x^2 = 0$, $x \varepsilon J$. Thus J is contained in the radical of $H_{2^{i+1}}$ and therefore $J = (0)$. Set $V = T^{2^{i+1}}$. By Lemma 4. 1, $V(S) = S$. The period of V is $2^{n-i-1} = r$. Arguing as in Lemma 4. 1 (3) and (4) and using P_r we have that

$$S = S \cap H_{2^{i+1}} \oplus S \cap K_{2^{i+1}} = S \cap K_{2^{i+1}}.$$

Thus $S \subset K_{2^{i+1}}$. However $K_{2^n} = (0)$ whence T is continuous.

4. 7. THEOREM. *Let B be a semi-simple Banach algebra. Suppose that every automorphism and anti-automorphism with period two on each closed semi-simple Banach sub-algebra B_1 of B is continuous, at least for B_1 where there is a continuous projection of B on B_1. Then every T of period 2^n on B is continuous, $n = 1, 2, \cdots$.*

The proof is by induction. The case $n = 1$ is clear. Suppose that the result holds for n and that T has period 2^{n+1}. Now by assumption, T^2 is continuous. Thus H_2 and K_2 are closed. Since $B = H_2 \oplus K_2$ it follows from [8], Lemma 1. 1. 1, that there is a continuous projection of B on H_2. H_2 is semi-simple by Lemma 4. 5. Thus T is continuous on H_2 since T has period one or two on H_2. Since $H_2 = H_1 \oplus H_2 \cap K_1$ and $H_2 \cap K_1 = \{x \varepsilon B \,|\, T(x) = -x\}$ by Lemma 2. 2, then $H_2 \cap K_1$ is closed. Also from Lemma 2. 2, $K_1 = K_2 \oplus H_2 \cap K_1$. By [7], p. 220, K_2 and $H_2 \cap K_1$ are disjoint closed linear manifolds and K_1 is closed. Then $S \subset K_1$ in view of Lemma 4. 1. The continuity of T now follows from Theorem 4. 6.

4. 8. COROLLARY. *Let B satisfy the conditions of Theorem 4. 7. Suppose*

that every automorphism of period k defined on B is continuous where k is an odd number. Then every T defined on B with period $2^n k$ is continuous, $n = 0, 1, 2, \cdots$.

For $n = 0$ we consider T an anti-automorphism of period k. Since k is odd it follows readily that such T can exist only if B is commutative. Then the continuity of T follows from [10], Corollary 6. 3.

Let T have period $2k$. Then T^2 is a continuous automorphism with period k. Let $k = 2r + 1$. T^k is continuous by Theorem 4. 7. Then so is $T = T^k T^{-2r}$.

Suppose that every T of period $2^{n-1}k$ is continuous where $n > 1$. Let T have period $2^n k$. Then T^2 is continuous. Arguing as above, we obtain the continuity of T. The result now follows by induction.

4. 9. LEMMA. *Let T have period three where H_1 and K_1 are closed and K_1 is a sub-algebra of B. Then S is a zero algebra.*

Since $B = H_1 \oplus K_1$ and H_1, K_1 are closed then ([7], p. 220) there exist constants $c_1, c_2 > 0$ such that, for $x = u + v$, $u \varepsilon H_1$, $v \varepsilon K_1$, $\| u \| \leqq c_1 \| x \|$ and $\| v \| \leqq c_2 \| x \|$. Let $x_i \varepsilon S$, $x_n^{(i)} \rightarrow x_i$, $x_n^{(i)} = u_n^{(i)} + v_n^{(i)}$, $u_n^{(i)} \varepsilon H_1$, $v_n^{(i)} \varepsilon K_1$ and $u_n^{(i)} + T(v_n^{(i)}) \rightarrow 0$, $i = 1, 2$. By Lemma 4. 1, $x_i \varepsilon K_1$. Also $\| v_n^{(i)} - x_i \| \leqq c_2 \| u_n^{(i)} + (v_n^{(i)} - x_i) \| \rightarrow 0$. Then $u_n^{(i)} \rightarrow 0$, $v_n^{(i)} \rightarrow x_i$ and $T(v_n^{(i)}) \rightarrow 0$. Since $I + T + T^2 = 0$ on K_1 and T^2 is an automorphism.

$$T^2(v_k^{(1)} v_k^{(2)}) = - v_k^{(1)} v_k^{(2)} - T(v_k^{(1)} v_k^{(2)}),$$

$$T^2(v_k^{(1)} v_k^{(2)}) = [v_k^{(1)} + T(v_k^{(1)})][v_k^{(2)} + T(v_k^{(2)})].$$

This gives

$$- 2 v_k^{(1)} v_k^{(2)} - T(v_k^{(1)} v_k^{(2)}) - T(v_k^{(1)}) T(v_k^{(2)}) = v_k^{(1)} T(v_k^{(2)}) + T(v_k^{(1)}) v_k^{(2)}.$$

Whether T is an automorphism or an anti-automorphism we obtain in the limit that $x_1 x_2 = 0$.

4. 10. THEOREM. *Let B be semi-simple. Then any T defined on B with period three where $S \subset K_1$ is continuous.*

It is sufficient to show in view of Theorem 2. 4 that T is continuous on S. By Lemmas 4. 1 and 4. 9, the separating ideal S_1 of T on S is a zero algebra and is, furthermore, by Theorem 2. 4, a two sided ideal of B. Thus $S_1 = (0)$ and T is continuous.

For the notion of spectrum in a real Banach algebra see [6] and [10]. Here $\rho(x)$ denotes the spectral radius of $x = \sup | \alpha |$, for α in the spectrum of x.

4.11. Lemma. *Let T have period two. Then for each $y \varepsilon H_1' \cap K_1$ $(K_1' \cap H_1)$ there exists a sequence $\{z_k\}$ in $H_1(K_1)$ such that $z_k \rightarrow y$, $\rho(z_k) \rightarrow 0$.*

Let $y \varepsilon H_1' \cap K_1 (K_1' \cap H_1)$ and let $\{x_k\}$ be a sequence in $H_1(K_1)$ such that $x_k \rightarrow y$. Set $w_k = (x_k + y)/2$. Then $w_k \rightarrow y$ and $w_k \varepsilon \bar{H}_1(\bar{K}_1)$. By a theorem of Hille, [3], Theorem 22.9.1, for each integer k there exists an element z_k in $H_1(K_1)$ such that simultaneously $\| w_k - z_k \| < k^{-1}$, $0 \leqq \rho(z_k) < \rho(w_k) + k^{-1}$. Note that

$$\rho(w_k) = \rho[T(w_k)] \leqq \| T(w_k) \| = \| x_k - y \|/2 \rightarrow 0.$$

Then we have $z_k \rightarrow y$ and $\rho(z_k) \rightarrow 0$.

4.12. Lemma. *Let T have period two on B. Suppose that $x_k \varepsilon H$, $k = 1, 2, \cdots, x_k \rightarrow x$, $x \varepsilon H$ and $\rho(x_k) \rightarrow 0$ imply that $x = 0$. Then K is closed. The condition on H is fulfilled if there exists a sub-additive function $| x |$ defined on H vanishing only for $x = 0$ such that $0 \leqq | x | \leqq \rho(x)$, $x \varepsilon H$.*

To show K closed it is sufficient to show that $K' \cap H = (0)$. Let $y \varepsilon K' \cap H$. By Lemma 4.11 there exists a sequence $\{z_k\}$ in K such that $z_k \rightarrow y$ and $\rho(z_k) \rightarrow 0$. Then $z_k^2 \varepsilon H$, $z_k^2 \rightarrow y^2 \varepsilon H$ and $\rho(z_k^2) \rightarrow 0$. By hypothesis this gives $y^2 = 0 = \rho(y)$. Applying the hypothesis to the sequence $\{y, y, \cdots\}$ we obtain $y = 0$.

For the second statement note that if $x_k \varepsilon H$, $x_k \rightarrow x$, $x \varepsilon H$ and $\rho(x_k) \rightarrow 0$ then

$$| x | \leqq | x - x_k | + | x_k | \leqq | x - x_k | + \rho(x_k) \rightarrow 0.$$

This gives $x = 0$.

If T is the involution in an A^*-algebra [10] or in a ρ^*-algebra [11], then such a sub-additive function exists on H.

The next result gives conditions under which all repeated square roots of a continuous automorphism are continuous.

4.13. Theorem. *Let B be semi-simple and let T have period two satisfying the condition of Lemma 4.12. Then any automorphism or anti-automorphism U of B such that $U^{2^n} = T$ is continuous, $n = 0, 1, 2, \cdots$.*

The case $n = 0$ is disposed of by Lemma 4.12 and Theorem 4.6. Suppose that the result holds for $n - 1$. Let $U^{2^n} = T$. Then U^2 is continuous. This gives the closure of H_2 and K_2 computed for U. Moreover H_2 is semi-simple by Lemma 4.5. We show that U is continuous on H_2. This is clear if $H_1 = H_2$; otherwise U has period two on H_2 (see arguments above, e.g. in Lemma 4.5). Since $H_1 \subset H$ (of T) the requirements of Lemma 4.12 are fulfilled by U on the Banach algebra H_2. By the above U is continuous on H_2. Then $H_2 \cap \mathrm{K}_1$ is closed. Arguing as in Theorem 4. 7, we see that K_1 is closed. Therefore U is continuous by Theorem 4. 6.

UNIVERSITY OF OREGON.

BIBLIOGRAPHY.

[1] S. Banach, *Théorie des opérations linéaires*, Monografje Matemacyzme, Warsaw (1932).

[2] B. Brown and N. H. McCoy, "Radicals and subdirect sums," *American Journal of Mathematics*, vol. 69 (1947), pp. 46-58.

[3] E. Hille, *Functional analysis and semi-groups*, American Mathematical Society Colloquium Publications, vol. 31, New York (1948).

[4] L. K. Hua, "On the automorphisms of a field," *Proceedings of the National Academy of Sciences*, vol. 35 (1949), pp. 386-389.

[5] N. Jacobson, "The radical and semi-simplicity for arbitrary rings," *American Journal of Mathematics*, vol. 67 (1945), pp. 300-320.

[6] I. Kaplansky, "Normed algebras," *Duke Mathematical Journal*, vol. 16 (1949), pp. 399-418.

[7] E. R. Lorch, "On a calculus of operators in reflexive vector spaces," *Transactions of the American Mathematical Society*, vol. 45 (1939), pp. 399-418.

[8] F. J. Murray, "On complementary manifolds and projections in the spaces L_p and l_p," *ibid.*, vol. 41 (1937), pp. 138-152.

[9] C. E. Rickart, "The singular elements of a Banach algebra," *Duke Mathematical Journal*, vol. 14 (1947), pp. 1063-1077.

[10] ———, "The uniqueness of norm problem for Banach algebras," *Annals of Mathematics*, vol. 51 (1950), pp. 615-628.

[11] B. Yood, "Topological properties of homomorphisms between Banach algebras," *American Journal of Mathematics*, vol. 76 (1954), pp. 155-167.

DE RHAM'S THEOREM FOR ARBITRARY SPACES.*

By J. SCHWARTZ.

Introduction. The abstract axiomatic approach to homology of Eilenberg and Steenrod (see Eilenberg-Steenrod [2]) yields such powerful uniqueness theorems that most of the principal isomorphism theorems of homology and cohomology theory can be derived from the Eilenberg-Steenrod theorems as corollaries. One notable exception has been the isomorphism between topological cohomology (say of Čech type) and the special cohomology for differentiable manifolds which can be defined in terms of forms. The difficulty has been simply this: the Eilenberg-Steenrod method requires that cohomology groups be defined for all triangulable spaces, and not every triangulable space is a differentiable manifold. How to define forms on more general spaces than manifolds? Taking our clue from a fundamental property of differentiable manifolds, we make use of *imbeddings in Euclidean spaces.*

The definition of the appropriate groups of forms is given in Section 1, together with the quite simple proofs of the seven cohomology axioms of Eilenberg-Steenrod. The groups can be defined for arbitrary compact subsets of Euclidean space; i. e., we get a cohomology theory for arbitrary compact finite dimensional spaces. In Section 2, we obtain the relationship between our "extrinsic" groups of forms and the ordinary "intrinsic" groups of forms in the case that our space is a differentiable manifold. In Section 3 we obtain a generalization of de Rham's classical theorem in its original form as a theorem of duality. In Section 4 we discuss the case of non-triangulable spaces, showing that our groups are simply the Čech cohomology groups with real coefficients. In Section 5 we discuss the theory of open manifolds.

§1. Definition and Basic Properties of the de Rham Cohomology Groups. 1. Pairs of spaces imbedded in a Manifold.

Let M be a C^∞ manifold (open or closed). Let S be a compact subset of M, and T a compact subset of S; so that (S, T) is a compact pair. We define spaces of forms as follows:

(a) $F_k(M)$ is the space of all C^∞ k-forms defined on M;

(b) $T_k(M; S)$ is the space of all C^∞ k-forms in $F_k(M)$ which vanish in a neighborhood of S;

---

* Received September 14, 1953.

(c) $Z_k(M; S, T) = \{f \varepsilon T_k(M; T) \mid df \varepsilon T_{k+1}(M; S)\}$;

(d) $F_k(M; S, T) = \{df + t \mid f \varepsilon T_{k-1}(M; T), t \varepsilon T_k(M, S)\}$.

We then put

$$H_k(M; S, T) = Z_k(M; S, T)/F_k(M; S, T).$$

The vector-space $H_k(M; S, T)$ may be called the k-th de Rham cohomology group of the pair S, T relative to the manifold M. We adopt the usual convention of taking all spaces to be zero-spaces for subscript $k < 0$. In this way, F_k, Z_k, T_k, and H_k are defined for all integer indices.

We shall show below that $H_k(M; S, T)$ is actually independent of M. For this reason, we begin our investigation by choosing some very large integer N, and supposing that all our spaces S, T are compact subsets of Euclidean space E^N. Wherever this is the case we will omit any specific notation for M in our notation. Thus we write $F_k(E^N)$ as F_k, $H_k(E^N; S, T)$ as $H_k(S, T)$, etc.

2. Mappings. Let $h: (S, T) \to (S_1, T_1)$ be a continuous mapping. Extend h to a continuous map $h': E^N \to E^N$. Let h_i be a sequence of C^∞ maps $E^N \to E^N$ which converge uniformly to h' on any compact subset of E^N. Then

LEMMA 1. *Let* $f \varepsilon Z_k(S_1, T_1)$. *Then, for sufficiently high* i, (a) $h_i^*f \varepsilon Z_k(S, T)$; (b) *the coset* mod $F_k(S, T)$ *of* h_i^*f *is independent of* i. *We call this coset the limiting coset of* h_i^*f.

COROLLARY. *The limiting coset of* h_i^*f *is independent of the approximating sequence* h_i.

LEMMA 2. *The limiting coset of* h_i^*f *is independent of the particular extension* h' *of* h *chosen.*

The proofs of these lemmas are to be found in the appendix to this section.

LEMMA 3. *If* $f \varepsilon F_k(S_1, T_1)$, *then the limiting coset of* h_i^*f *is zero.*

Proof. $f = dg + t$, $g \varepsilon T_{k-1}(T_1)$ and $t \varepsilon T_k(S_1)$. It is clear from the definition that $h_i^*g \varepsilon T_{k-1}(T)$ and $h_i^*t \varepsilon T_k(S)$ for all sufficiently large i. This shows, however, that $h_i^*f = dh_i^*g + h_i^*t$ is in $F_k(S, T)$ for sufficiently large i, so that the limiting coset of h_i^*f is zero.

Definition 1. If $g \varepsilon H_k(S_1, T_1)$, and z is an element of the coset g, we put $h^*g =$ limiting coset of h_i^*z. Lemmas 1, 2, and 3 ensure that this is a single-valued definition. It is clear that $h^*: H_k(S_1, T_1) \to H_k(S, T)$ is a linear map.

LEMMA 4. *If* $h: (S, T) \to (S_1, T_1)$, *and* $h_1: (S_1, T_1) \to (S_2, T_2)$, *then* $(h_1 h)^* = h^* h_1^*$.

Proof. Let h' and h_1' be extensions of h and h_1 to all of E^N. Let h_i and $h_i^{(1)}$ be sequences of C^∞ maps which converge uniformly on any compact set to h' and h_1', respectively. Then $h_i^{(1)} h_i$ is a sequence of C^∞ maps which converge uniformly on any compact set to $h_1' h'$, which is an extension of $h_1 h$. If $g \varepsilon H_k(S_2, T_2)$, and z is an element of the coset g, $(h_1 h)^* g =$ limiting coset of $(h_i^{(1)} h_i)^* z =$ limiting coset of $h_i^* (h_i^{(1)} z)$. From the fact that the limiting coset of $h_i^{(1)} z$ is $h_1^* z$, and from Lemma 3, the equation $(h_1 h)^* g = h^* h_1 g$ follows readily.

LEMMA 5. *If i is the identity map* $i: (S, T) \to (S, T)$, *then* $i^*: H_n(S, T) \to H_n(S, T)$ *is the identity map.*

Proof. As an extension of i we take i', the identity map of E^N. Then we can approximate i' by $h_j = i'$; it follows that each h_j^* is an identity map, so i^* is also an identity map.

LEMMA 6. *Suppose that U is a relatively open subset of S, and that the closure of U and the closure of $S - T$ are disjoint, i. e.* $\mathrm{cl}(U) \cap \mathrm{cl}(S - T) = \emptyset$. *Let i be the natural identity map of* $(S - U, T - U) \to (S, T)$. *Then*

$$i^*: H_n(S, T) \to H_n(S - U, T - U)$$

is an isomorphism onto.

Proof. (A) i^* is onto. Let $z \varepsilon Z_k(S - U, T - U)$. Since $\mathrm{cl}(U) \cap \mathrm{cl}(S - T) = \emptyset$, we can find a C^∞ function f defined on all of E^n, such that $f \equiv 1$ on a neighborhood of $\mathrm{cl}(S - T)$; $f \equiv 0$ on $\mathrm{cl}(U)$. Then (1) $f \wedge z = 0$ near $\mathrm{cl}(U)$ and near $T - U$, i. e., $f \wedge z \varepsilon T_k(T)$. (2) $d(f \wedge z) = df \wedge z + f \wedge dz$. We have $dz = 0$ near $S - U$, $f = 0$ near $\mathrm{cl}(U)$, so the second term of this expression is in $T_{k+1}(S)$. We have $df = 0$ near $\mathrm{cl}(U) \cap \mathrm{cl}(S - T)$, and $z = 0$ near $T - U$, so that the first term $df \wedge z$ is in $T_{k+1}(S)$. Taking all together, we see that $f \wedge z \varepsilon Z_k(S, T)$. (3) Write $z' = f \wedge z$. Then $z - z' = (1 - f) \wedge z$. Since $1 - f \equiv 0$ near $\mathrm{cl}(S - T)$, and $z \equiv 0$ near $T - U$, we have $z - z' \varepsilon T_k(S - U)$, so $z - z' \varepsilon F_k(S - U, T - U)$. Thus, z' and z belong to the same class ν in $H_k(S - U, T - U)$. If ν' is the class in $H_k(S, T)$ which contains z', it is clear that $i^*\nu' = \nu$. Since z was an arbitrary element of $Z_k(S - U, T - U)$, we have proved i^* is onto.

(B) i^* is one-to-one. For, let $\nu \varepsilon H_k(S, T)$, and suppose that $i^*\nu = 0$. Then, if $z \varepsilon Z_k(S, T)$ belongs to the class ν, it follows that

$$z = dg + t, \qquad g \varepsilon T_{k-1}(T - U), \qquad t \varepsilon T_k(S - U).$$

Since $z \equiv 0$ near T, it follows that $dg = -t$ near T. We then choose a C^∞ function f as in part (A) of the proof of this lemma, and write

$$z = d(f \wedge g) + (1 - f) \wedge dg - df \wedge g + f \wedge t + (1 - f) \wedge t;$$

since $f = 0$ near $\mathrm{cl}(U)$, and $g = 0$ near $T - U$, $f \wedge g \varepsilon T_{k-1}(T)$, and we have only to show that

$$m = (1 - f) \wedge dg - df \wedge g + f \wedge t + (1 - f) \wedge t \,\varepsilon\, T_k(S)$$

to prove that $z \varepsilon F_k(S, T)$, so that $\nu = 0$. Now, near T, m is $f \wedge t - df \wedge g$. We have $df = 0$ near $\mathrm{cl}(U)$, $g = 0$ near $T - U$; $f = 0$ near U, $t = 0$ near $S - U$; hence $m = 0$ near T. Near $\mathrm{cl}(S - T)$ we have $(1 - f) \equiv 0$, so that $m = f \wedge t - df \wedge g$ once more. Since t and df are zero near $\mathrm{cl}(S - T)$, $m = 0$ near $\mathrm{cl}(S - T)$. Altogether $m = 0$ near S, so that $m \varepsilon T_k(S)$.

3. The boundary operator. From the definitions of the various spaces it is clear that (a) $d : F_k(T) \rightarrow K_{k+1}(S, T)$, (b) $d : Z_k(T) \rightarrow Z_{k+1}(S, T)$. Thus, d defines a natural map $d^* : H_k(T) \rightarrow H_{k+1}(S, T)$ by passage to factor groups.

Once d^* is defined, we can follow Eilenberg-Steenrod by letting $j : T \rightarrow S$ and $i : S \rightarrow (S, T)$ be the natural identity maps, and considering the "homology" sequence of groups and homomorphisms

$$\cdots \xleftarrow{i^*} H_{k+1}(S, T) \xleftarrow{d^*} H_k(T) \xleftarrow{j^*} H_k(S) \xleftarrow{i^*} H_k(S, T) \xleftarrow{d^*} \cdots.$$

LEMMA 7. *The homology sequence is exact.*

Proof. (a_1) If ν is an element of $H_k(T)$, and $z \varepsilon Z_k(T)$ is in the class ν, then $i^*d^*\nu$ is evidently the class in $H_{k+1}(S)$ of dz. Since $dz \varepsilon F_{k+1}(S)$, $i^*d^*\nu = 0$.

(b_1) If ν is an element of $H_k(T)$, and $z \varepsilon Z_k(S)$ is in the class ν, then $d^*j^*\nu$ is evidently the class in $H_{k+1}(S, T)$ of dz. Since $dz \varepsilon T_{k+1}(S)$, $d^*j^*\nu = 0$.

(c_1) If ν is an element of $H_k(S, T)$, and $z \varepsilon Z_k(S, T)$ is in the class ν, then $j^*i^*\nu$ is evidently the class in $H_k(T)$ of z. Since $z \varepsilon T_k(T) \subseteq F_k(T)$, it follows that $j^*i^*\nu = 0$.

(a_2) Let $\nu \in H_k(T)$, $d^*\nu = 0$. If $z \in Z_k(T)$ is in the class ν, it follows that $dz \in F_{k+1}(S, T)$, so that $dz = dz_1 + z_2$, $z_1 \in T_k(T)$, $z_2 \in T_{k+1}(S)$. I. e., $d(z - z_1) \in T_{k+1}(S)$; $z - z_1 \in Z_k(S)$. Thus the class $\nu' \in H_k(T)$ of $z - z_1$ is in $j^*H_k(S)$. Since $z_1 \in T_k(T) \subseteq F_k(T)$, $\nu = \nu'$.

(b_2) Let $\nu \in H_k(S)$, $j^*\nu = 0$. If $z \in Z_k(S)$ is in the class ν, it follows that $z \in F_k(T)$; i. e., $z = dz_1 + z_2$, $z_1 \in F_k$, $z_2 \in T_k(T)$. Thus $z - dz_1 \in T_k(T)$; so that $z - dz_1 \in Z_k(S, T)$. If ν' is the class in $H_k(S)$ of $z - dz_1$, it follows that $\nu' \in i^*H_k(S, T)$. But, since $dz_1 \in F_k(S)$, $\nu' = \nu$.

(c_2) Let $\nu \in H_k(S, T)$, $i^*\nu = 0$. If $z \in Z_k(S, T)$ is an element of the class ν, it follows that $z \in F_k(S)$; i. e., $z = dz_1 + z_2$, $z_1 \in F_{k-1}$, $z_2 \in T_k(S)$. Then $z - z_2 = dz_1$; since $z - z_2 \in T_k(T)$, $z_1 \in Z_{k-1}(T)$. It follows that the class ν' of $z - z_2$ in $H_k(S, T)$ is in $d^*H_{k-1}(T)$; since $z_2 \in T_k(T) \subseteq F_k(T)$, $\nu' = \nu$.

LEMMA 8. *Let* $h\colon (S, T) \to (S_1, T_1)$ *be a continuous map, and let* $h_1\colon T \to T_1$ *be its restriction to* T. *If* $\nu \in H_{k-1}(T_1)$, *then* $d^*h_1^*\nu = h^*d^*\nu$.

Proof. Let h' be an extension of h to all of E^N, and let h_i be a sequence of C^∞ maps of $E^N \to E^N$ which converge uniformly to h' on every compact set. Then, if $z \in Z_{k-1}(T)$, $h^*d^*\nu =$ limiting coset in $H_k(S, T)$ of h_i^*dz $=$ limiting coset in $H_k(S, T)$ of dh_i^*z. Since the limiting coset of h_i^*z is $h_1^*\nu$, our lemma is now evident.

LEMMA 9. *Let* $h_0, h_1\colon (S, T) \to (S_1, T_1)$ *be continuons maps which are homotopic. Then* $h_0^* = h_1^*$.

LEMMA 10. *Let* z *be a* k*-form such that* $dz \equiv 0$ *near a point* p, $k > 1$. *Then we can find a form* z_1 *such that* $z \equiv dz_1$ *near* p.

These lemmas are proved in the appendix to this section. Lemma 10 evidently implies: If P is a one point space, $H_k(P) = 0$ for $k > 0$. In dimension 0, $Z_0(P)$ is the space of all scalar functions constant near P, and $F_0(P) = T_0(P)$ is the space of all scalar functions zero near P. Thus $H_0(P)$ is isomorphic with the field of real numbers.

We have now verified all properties of Eilenberg-Steenrod, and can conclude that we have

THEOREM 1. *For every pair* (S, T) *of triangulable spaces,* $H_k(S, T)$ *is naturally isomorphic to the relative* k*-th cohomology group of* (S, T), *no matter how these cohomology groups are defined.*

4. Appendix. Homotopy properties of forms.

Lemma 11. *Let M and N be C^∞ manifolds, and let f be a k-form on N. Suppose that $M \supseteq S \supseteq T$, where S and T are compact, and that we can define a map $h_t : M \to N$ for all $-1 < t < 2$ such that $h_t(m)$ is a C^∞ function both of the real variable t and of $m \in M$. If, for $0 \leqq t \leqq 1$,*

(a) $h_t(S) \cap \operatorname{car}(df) = \varnothing$, (b) $h_t(T) \cap \operatorname{car}(f) = \varnothing$,

then $h_0{}^*f$ and $h_1{}^*f$ are in the same coset of $Z_k(M; S, T)$ modulo $F_k(M; S, T)$.

Before giving the proof, we show how Lemma 11 implies Lemmas 1, 2, 9, and 10.

Proof of Lemma 1. Put $h_t{}^{(i,j)}(x) = th_i(x) + (1-t)h_j(x)$, and apply Lemma 11.

Proof of Lemma 2. Let h' and g' be two extensions of h, and let h_i approach h, g_i approach g, both uniformly on every compact set, where h_i and g_i are C^∞ maps. Put $h_t{}^{(i)}(x) = th_i(x) + (1-t)g_i(x)$, and apply Lemma 11.

Proof of Lemma 9. Let $h_t : (S, T) \to (S_1, T_1)$ be a continuous homotopy between h_0 and h_1. Extend h_t to a map $h_t' : E^N \to E^N$ which is defined and continuous for $-\infty < t < +\infty$, $x \in E^N$. Let $h_t{}^{(i)} : E^N \to E^N$ be a sequence of C^∞ maps depending in a C^∞ manner on the real parameter t, and such that $h_t{}^{(i)}$ approaches h_t' uniformly on every compact subset of E_n, uniformly for every bounded set of parameter values t. Let $z \in Z_k(S_1, T_1)$. Then $h_0{}^*z =$ limiting coset of $h_0{}^{(i)*}z$, $h_1{}^*z =$ limiting coset of $h_1{}^{(i)*}z$. Since, by Lemma 11, coset of $h_1{}^{(i)*}z =$ coset of $h_2{}^{(i)*}z$ for all sufficiently large i, Lemma 9 is proved.

Proof of Lemma 10. Put $h_t(x) = t(x-p) + p$. Then $h_1{}^*z = z$ and $h_0{}^*z = 0$ for every k-form z, $k > 0$. If $z \in Z_k(p)$, it follows by Lemma 11 that $z \in F_k(p)$; so from this, Lemma 10 is evident.

Proof of Lemma 11. We let I be the open interval $-1 < t < 2$, and let M' be the direct-product manifold $M \times I$. Define $H : M' \to N$ by $H[m, t] = h_t(m)$. Define $p : M' \to M$ by $p[m, t] = m$. Then inspection of the definitions reveals

$$(H^*f)[m, t] = p^*\{(h_t{}^*f)(m)\} + (p^*\beta_t{}^{(m)}) \wedge dt,$$

where β_t is a C^∞ $k-1$ form on M which depends in a C^∞ fashion on the real parameter t. From $h_t(T) \cap \operatorname{car}(f) = \varnothing$ it follows readily that $\operatorname{car}(\beta_t(m)) \cap T = \varnothing$. We have then

$$(H^*df)[m,t] = \pm\, \partial/\partial t p^*\{(h_t^*)(m)\} \wedge dt \pm p^*\{(h_t^*df)(m)\} + p^*d\beta_t \wedge dt.$$

Since $h_t(S) \cap \operatorname{car}(df) = 0$, $H^*df = 0$ near $S \times [0,1]$. Likewise, $h_t^*df = 0$ near $S \times [0,1]$. It follows that $\partial p^*(h_t^*f)/\partial t = \pm\, p^*(d\beta_t)$ or integrating from 0 to 1,

$$p^*(h_0^*f - h_1^*f \pm d\int_0^1 \beta_t dt) = 0; \text{ hence } h_0^*f - h_1^*f = \pm\, d\int_0^1 \beta_t dt.$$

Put $\pm \int_0^1 \beta_t dt = g$. Then $g \in T_{k-1}(T)$, so that $h_0^*f - h_1^*f \in F_k(S,T)$.

§2. Intrinsic Forms on a Manifold.

Definition 2. Let M be a C^∞ manifold. A C^∞ map $h: M \to E^N$ which determines a one-to-one map of the tangent vector space at every point on M is called a regular imbedding of M in E^N.

Definition 3. If, in Definition 2, M is a subset of E^N and the natural identity map $i: M \to E^N$ is a regular imbedding, we say that M is regularly imbedded in E^N.

LEMMA 12. *If M is a C^∞ manifold regularly imbedded in E^N, then there exists a neighborhood U and a map $n: U \to M$ such that the map $n: U \to M$ is a C^∞ retraction of U on M.*

Proof. We consider the C^∞ N-dimensional manifold Σ of segments $\sigma = [a,b]$ such that $a \in M$, and σ is orthogonal to M at its foot a. It is clear that the map $t: \Sigma \to E^N$ determined by $t[a,b] = b$ is a C^∞ map whose Jacobian determinant is non-zero at every point σ in the set of points Σ_0 of zero length. It follows by the implicit function theorem that there exists a neighborhood Φ of Σ_0 such that the map $t: \Phi \to E^N$ is C^∞, one-to-one, and maps onto an open set in E^N; t^{-1} is also C^∞ by the implicit function theorem. We can now put $s[a,b] = a$ and $n = st^{-1}$. Q. E. D.

Suppose that h is a regular imbedding of M in E^N. Put $\bar{S} = h(S)$, $\bar{T} = h(T)$. Then

$$h^*: Z_k(\bar{S},\bar{T}) \to Z_k(M;S,T), \qquad h^*: F_k(\bar{S},\bar{T}) \to F_k(M;S,T).$$

By passage to cosets we define, in the natural way, a map

$$h^*: H_k(\bar{S},\bar{T}) \to H_k(M;S,T).$$

THEOREM 2. *If h is*

(a) *A regular imbedding of the manifold M,*

(b) *A non-singular linear map of the Euclidean space M, then h^* is an isomorphism onto.*

Proof. Without loss of generality, we can assume that M is regularly imbedded in E^N, and that h is the identity map of $M \rightarrow E^N$. Let n be the map whose existence is established in Lemma 12; $n: U \rightarrow M$, where U is a neighborhood of M. Suppose that $z \,\varepsilon\, Z_k(M; S, T)$. Then n^*z is defined and C^∞ in U. Let z_1 be a C^∞ form defined in all E^N such that $z_1 \equiv n^*z$ in a neighborhood of S. Then $z_1 \,\varepsilon\, Z_k(S, T)$, and $h^*z_1 = h^*n^*z = (nh)^*z = i^*z = z$ in a neighborhood of S. Thus, if $\nu \,\varepsilon\, H_k(M; S, T)$ is the coset of z, and $\nu' \,\varepsilon\, H_k(S, T)$ that of z_1, $\nu = h^*\nu'$; so that h^* is a mapping onto.

Conversely, suppose that $\nu' \,\varepsilon\, H_k(S, T)$, and that $h^*\nu' = 0$. If $z' \,\varepsilon\, Z_k(S, T)$ is an element of the class ν', it follows that $h^*z' = dg_1 + g_2$, $g_1 \,\varepsilon\, T_{k-1}(M; T)$, $g_2 \,\varepsilon\, T_k(M; S)$. Let g_1' and g_2'; z'' be C^∞ forms defined on all of E^N such that $g_1' = n^*g_1$ near S, $g_2' = n^*g_2$ near S, $z'' = n^*h^*z'$ near S. Then $z'' = dg_1' + g_2' + g_3'$, where $g_1' \,\varepsilon\, T_k(T)$, $g_2' \,\varepsilon\, T_k(S)$, $g_3' \,\varepsilon\, T_k(S)$; so $z'' \,\varepsilon\, F_k(S, T)$. It only remains to show that $z'' - z' \,\varepsilon\, F_k(S, T)$, and this we do as follows:

Let $S_\epsilon = \{x \,\varepsilon\, E^N \mid |x - s| < \epsilon \text{ for some } s \,\varepsilon\, S\}$. The set U contains some set S_{ϵ_0}. Let $\eta = \frac{1}{2}\epsilon_0$, and put $H_t(m) = tm + (1 - t)n(m)$, so that $H_t: S_\eta \rightarrow S_{\epsilon_0}$ for $-1 < t < 2$. The open sets S_{ϵ_0} and S_η can be regarded as manifolds. By Lemma 11, $H_0{}^*z' - H_1{}^*z' \,\varepsilon\, F_k(S_\eta; S, T)$. Now, it is easily seen that $H_0{}^*z' = n^*h^*z' = z''$ near S, and that $H_1{}^*z' = z$ near S. Thus $z'' - z' - dz_1 = 0$ near S, where z_1 is a suitably chosen form in $T_{k-1}(S_\eta; S, T)$. If z_1' is defined in all of E^N in such a way as to ensure $z_1 = z_1'$ near S, we have $z'' - z' - dz_1' = 0$ near S, so that $z'' - z' \,\varepsilon\, F_k(S, T)$.

§3. De Rham's theorem.

LEMMA 13. *Let H and $\mathfrak{h}$ be a cohomology and homology theory with coefficient-field equal to the field R of real numbers, both defined on an admissible category containing all triangulable pairs and their maps. Suppose that a bilinear multiplication (x, y) exists between elements of $H_n(A, B)$ and $\mathfrak{h}^n(A, B)$ such that*

(1) $(x, d_*y) = (d^*x, y)$ *for* $x \,\varepsilon\, H_{n-1}(B)$, $y \,\varepsilon\, \mathfrak{h}^n(A, B)$.

(2) $(x, j_*y) = (j^*x, y)$ *for* $x \,\varepsilon\, H_n(A_1, B_1)$, $y \,\varepsilon\, \mathfrak{h}^n(A, B)$, *and* $j: (A, B) \rightarrow (A_1, B_1)$ *an admissible map.*

(3) *The multiplication (x, y) between $H_0(P)$ and $H^0(P)$ is a dual pairing.*

Then the multiplication between $H_n(A,B)$ and $\mathfrak{h}^n(A,B)$ is a dual pairing for every triangulable pair.

Proof. We make use of Theorem III. 10. 1 of Eilenberg-Steenrod [2]. We introduce the space $H^n(A,B)$ of all linear functionals on $H_n(A,B)$, and let d_* and j_* be the linear maps of the spaces $H^n(A,B)$ dual to the corresponding maps d^*, j^* of $H_n(A,B)$. Then it is clear that the groups $H(A,B)$, and the maps d_* and j_* determine a homology theory. Now there exists a unique linear mapping $h: H^n(A,B) \rightarrow \mathfrak{h}^n(A,B)$ such that $(x, hy) = y(x)$ for $y \varepsilon H^n(A,B)$. In virtue of (1) and (2), this map has properties (2) and (3) of the Eilenberg-Steenrod Theorem III. 10. 1. In virtue of (3), the map $h: H^0(P) \rightarrow \mathfrak{h}^0(P)$ is an isomorphism onto. It then follows by the Eilenberg-Steenrod Theorem III. 10. 1 that $h: H^n(A,B) \rightarrow H^n(A,B)$ is an isomorphism onto for every triangulable pair. But this clearly means: the multiplication between $H_n(A,B)$ and $\mathfrak{h}^n(A,B)$ is a dual pairing for every triangulable pair. Q. E. D.

Let $\mathfrak{h}^n(A,B)$ denote the singular homology groups of the pair (A,B). In the particular case where (A,B) is a pair of open sets in Euclidean space, let $\mathfrak{h}_\infty{}^n(A,B)$ denote the relative singular groups based on C^∞ chains. Then there is a natural homomorphism $\tau: \mathfrak{h}_\infty{}^n(A,B) \rightarrow \mathfrak{h}^n(A,B)$. The map τ is an isomorphism onto. The absolute case of this result is well known; cf. Eilenberg [1]. The relative case results easily from the absolute by the use of the "five" lemma, Eilenberg-Steenrod [2], Lemma I. 4. 3. We can make use of these facts as follows. Let α be a relative singular cycle of a compact pair (A,B), where $A \subseteq E^N$. Let $z \varepsilon Z_n(A,B)$. Then we take a C^∞ cycle α_1 of $(E^N - \operatorname{car}(dz), E^N - \operatorname{car}(z)) = (U, V)$ which is in the same class of $h^n(U, V)$ as α. We put $(\alpha, z) = \int_{\alpha_1} z$. This multiplication is single-valued, since if α_2 is also a C^∞ cycle in the same class of $\mathfrak{h}^n(U, V)$ as α, then α_1 and α_2 are in the same class of $\mathfrak{h}_\infty{}^n(U, V)$, so that $\int_{\alpha_1} z = \int_{\alpha_2} z$ by Stoke's theorem. If α is in the zero class of $\mathfrak{h}^n(A,B)$, we may take $\alpha_1 = 0$, so that it follows that in this case, $(\alpha, z) = 0$. If $z \varepsilon F_n(A,B)$, then we can write $z = dz_1 + z_2$, $z_1 \varepsilon T_{n-1}(B)$, $z_2 \varepsilon T_{n-1}(A)$. We let $U_1 = E^N - \operatorname{car}(z_2)$; $V_1 = (E^N - \operatorname{car}(z_1)) \cap U_1$. Then we may suppose that α_1 is in the same class of $\mathfrak{h}^n(U_1, V_1)$ as α; in this case it follows by Stoke's theorem that

$$\int_{\alpha_1} (dz_1 + z_2) = \int_{\alpha_1} dz_1 = \int_{d\alpha_1} z_1 = 0.$$

Thus the multiplication (α, z) leads by passage to cosets to a multiplication

(h_1, h) between $\mathfrak{k}^n(A, B)$ and $H_n(A, B)$. Once we verify that this multiplication has the properties demanded in Lemma 13, we will have proved

THEOREM 3. *If (A, B) is a triangulable pair, then the multiplication between $\mathfrak{k}^n(A, B)$ and $H_n(A, B)$ introduced above is a dual pairing between these groups.*

If we now recall Theorem 2, and make use of the triangulation theorem for compact C^∞ manifolds and the isomorphism between the singular homology groups of such a manifold and the corresponding groups based on C^∞ chains (as given, say, in Eilenberg [1]), we can conclude

THEOREM 4. (*De Rham*). *Let S be a compact C^∞ manifold. Then* (a) *A closed form on S exists which has arbitrarily prescribed periods on independent cycles of S.*

(b) *If a closed form on S has period zero on every cycle of S, it is the exterior derivative of a form on S.*

To prove Theorems 3 and 4, we have, by Lemma 13, only to verify certain properties of the multiplication (h_1, h):

(1) Let $z \varepsilon Z_{n-1}(B)$, and let α be a singular cycle of A modulo B. Let $(U, V) = (E^N, E^N - \operatorname{car}(dz))$. Then if α_1 is a C^∞ cycle in the same class in $\mathfrak{k}^n(U, V)$ as α, $d\alpha_1$ is in the same class of $\mathfrak{k}^n(V)$ as $d\alpha$. Thus we have

$$(\alpha, dz) = \int_{\alpha_1} dz, \qquad (d\alpha, z) = \int_{d\alpha_1} z.$$

By Stoke's theorem, these two expressions are equal.

(2) Let $j: (A, B) \to (A_1, B_1)$. Let $z \varepsilon Z_n(A_1, B_1)$, and let $j': E^N \to E^N$ be an extension of j. Let j_i be a sequence of C^∞ maps which converge uniformly to j' on every compact set. Then for i sufficiently large, j_i^*z is in the same class of $H_n(A, B)$ as j^*z. Choose some such i. Let α be a singular cycle of A modulo B. We put

$$(U_1, V_1) = (E^N - \operatorname{car}(dz), E^N - \operatorname{car}(z)), \qquad (U, V) = (j_i)^{-1}(U_1, V_1).$$

We let (U_2, V_2) be a pair of neighborhoods of (A, B) such that $(U_2, V_2) \subseteq (U, V)$, and such that if $x \varepsilon U$ (or, V) then the whole segment between $j(x)$ and $j_i(x)$ is in U_1 (or, V_1). Let α_2 be a C^∞ cycle in the same class of $\mathfrak{k}^n(U_2, V_2)$ as α; then it follows by homotopy that $j_i\alpha_2$ is in the same class of $\mathfrak{k}^n(U_1, V_1)$ as $j\alpha$. Hence

$$(j\alpha, z) = (j_i\alpha_2, z) = (\alpha_2, j_i^*z) = (\alpha_2, j^*z).$$

(3) That $H_0(P)$ and $\mathfrak{K}^n(P)$ are dually paired is clear from the definitions.

§4. **The de Rham groups of general pairs.** All question as to the nature of the de Rham groups for non-triangulable pairs (S, T) is settled by the following

THEOREM 5. *The groups $H_n(S, T)$ are naturally isomorphic to the Čech cohomology groups (with real coefficients) of the compact pair S, T.*

Proof. Let (S_k, T_k) be a sequence of triangulable pairs such that $(S_k, T_k) \supseteq (S_{k+1}, T_{k+1})$ and $\bigcap_{k=1}^{\infty} (S_k, T_k) = (S, T)$. Then the identity map $i_k : (S_{k+1}, T_{k+1}) \to (S_k, T_k)$ determines a map $i_k^* H_l(S_k, T_k) \to H_l(S_{k+1}, T_{k+1})$, so that we have a direct sequence of homology groups. In the same way, the identity map $j_k : (S, T) \to (S_k, T_k)$ determines a map $j_k^* : H_l(S_k, T_k) \to H_l(S, T)$. It is clear from the definitions of the groups that each $\alpha \varepsilon H_l(S, T)$ is in $j_k^* H_l(S_k, T_k)$ for all sufficiently large k. Moreover, $j_k^* \beta = 0$ for $\beta \varepsilon H_l(S_k, T_k)$ if and only if $i_{k+p}^* i_{k+p-1}^* \cdots i_k^* \beta = 0$ for some sufficiently large p. It follows that $H_l(S, T) =$ direct limit $H_l(S_k, T_k)$, and once this is established, our result follows readily from Theorem X. 3. 1 (Cf. also Theorem X. 12. 1) of Eilenberg-Steenrod [2].

§5. **Open manifolds.** Suppos that M is a manifold which is not compact but only the union of a countable infinity of compact subsets, and that we consider subsets S, T which are not compact but only closed. Then there are two principal types of de Rham cohomology groups which can be introduced on M:

(a) If we let all forms be forms with compact carriers, we get spaces which can be denoted $F_k^c, Z_k^c, \cdots, H_k^c(M; S, T)$.

(b) If we let all forms be forms with unrestricted carriers, we get spaces of forms which can be denoted $F_k^u, Z_k^u, \cdots, H_k^u(M; S, T)$.

Of course, if S is compact, or if the closure of the complement of T is compact, the groups $H_k^u(M; S, T)$ and $H_k^c(M; S, T)$ are naturally isomorphic. In either of these two cases we can drop the superscript and write simply $H_k(M; S, T)$ as in §1-4.

The groups $H_k^c(M; S, T)$ seem to be the more interesting of the two types of de Rham groups. Part of this interest stems from the following obvious but important

THEOREM 6. *If M is a C^∞ manifold, and (S, T) is a pair of closed subsets of M, then $H_k{}^c(M; S, T)$ and $H_k{}^c(M - T; S - T)$ are naturally isomorphic.*

To prove this theorem, we have only to examine the definitions of the two groups. As is well known, Theorem 6 is essentially involved in the proof of separation theorems by means of homology theory.

The groups $H_k{}^c$ and $H_k{}^u$ are related to the two types of cohomology groups which we can obtain by taking direct and inverse limits over the directed family of compact subsets of M. Indeed, let C_n be a sequence of compact subsets of M such that interior $(C_{n+1}) \supseteq C_n$ and such that $M = \bigcup_{i=1}^{\infty} C_n$. Let D_n be the complement of interior (C_n). Let $S_n{}^u = C_nS$, $T_n{}^u = C_nT$. Let $S_n{}^c = SD_n$, $T_n{}^c = TD_n$. Then we have identity maps

$$[*] \qquad i_n : (S_n{}^u, T_n{}^u) \to (S_{n+1}{}^u, T_{n+1}{}^u),$$

$$[**] \qquad j_n : (S_{n+1}{}^c, T_{n+1}{}^c) \to (S_n{}^c, T_n{}^c).$$

These maps determine homomorphisms

$$i_n{}^* : H_k(M; S_{n+1}{}^u, T_{n+1}{}^u) \to H_k(M; S_n{}^u, T_n{}^u),$$

$$j_n{}^* : H_k(M; S_n{}^c, T_n{}^c) \to H_k(M; S_{n+1}{}^c, T_{n+1}{}^c).$$

Thus, the groups $H_k(M; S_n{}^u, T_n{}^u)$ form an inverse sequence, and the groups $H_k(M; S_n{}^c, T_n{}^c)$ form a direct sequence. We have

THEOREM 7c. *$H^c(M; S, T) =$ direct limit $H(M; S_n{}^c, T_n{}^c)$; such a general result is not true for the groups $H_k{}^u$, but we have the following weaker*

THEOREM 7u. *Let M be triangulated, and let S and T be subsomplexes of M. Then*

$$H_k{}^u(M; S, T) = \text{inverse limit } H_k(M; S_n{}^u, T_n{}^u).$$

Sketch of Proof of Theorem 7c. Let $\bar{H}_k(M; S, T) =$ direct limit $H_k(M; S_n{}^c, T_n{}^c)$. The elements of $\bar{H}_k(M; S, T)$ are certain equivalence classes of elements of $\bigcup_{n=1}^{\infty} H_k(M; S_n{}^c, T_n{}^c)$. (For the exact definition of the direct limit group, see Eilenberg-Steenrod [2], pp. 220-222.) Let $\zeta \varepsilon \bar{H}_k$, and let $\nu \varepsilon H_k(M; S_n{}^c, T_n{}^c)$ be in the equivalence class ζ. Let z be a closed form of the class ν. Then z is a closed form with compact carrier, and hence z determines an element η of the group $H_k{}^c(M; S, T)$ such that $z \varepsilon \eta$. We put $h\zeta = \eta$. It follows readily from the exact definition of the equivalence relation determining the class ζ that the map h is a single-valued linear isomorphism of $\bar{H}_k(M;S,T)$ onto $H_k^c(M;S,T)$.

A simple example suffices to show that Theorem 7*u* cannot be stated in as simple a form as Theorem 7*c*. Indeed, let M be the Euclidean plane. Let A_∞ be the union of the three segments connecting the four points $(\infty, 1)$, $(0, 1)$, $(0, -1)$, and $(\infty, -1)$ in the order indicated. Let A_n be the union of the three segments connecting the four points $(\frac{1}{n}, 1-\frac{1}{n})$, $(n, 1-\frac{1}{n})$, $(n, -1+\frac{1}{n})$, and $(\frac{1}{n}, -1+\frac{1}{n})$ in the order indicated. Let $S = A_\infty \cup \bigcup_{i=2}^{\infty} A_n$. Then S is a closed subset of the place. Let C_n be the closed square surface with horizontal sides, center at the origin, and side-length $2n+1$. It is clear that $H_1(S \cap C_n) = 0$, so that

$$\text{inverse limit } H_1(S \cap C_n) = 0.$$

On the other hand, we shall show that $H_1^n(M;S)$ is not zero. Let $\theta(P)$ be the angle that the arc from $(1,0)$ to p makes with the horizontal, and let $r(p)$ be the length of this arc. Let $v(p)$ be a C^∞ vector field in M which has the value $(r^{-1}\sin\theta, -r^{-1}\cos\theta)$ in the neighborhood of S. Then v is closed in the neighborhood of S. Now, if $v = df$ in the neighborhood U of S, then $\int_C v = 0$ for every closed arc in U. On the other hand, it is clear that every neighborhood of S contains a closed arc C surrounding the origin once in the positive sense, and we have $\int_C v = 2\pi$ along any such arc.

Let us now proceed to the

Proof of Theorem 7u. Let $\bar{H}_k(M;S,T) =$ inverse limit $H_k(M;S_n^u,T_n^u)$. We define a homomorphism h of $H_k^u = H_k^u(M;S,T)$ into $\bar{H}_k = \bar{H}_k(M;S,T)$ as follows: let $\alpha \varepsilon H_k^u$, and let z be a form in the class α. Then $z = 0$ in a neighborhood of T, and $dz = 0$ in a neighborhood of S. Hence $z \varepsilon Z_k(M;S_n^u,T_n^u)$, so that z belongs to a well-determined class $\alpha_n \varepsilon H_k(M;S_n^u,T_n^u)$. It is easily seen that $i_n^*\alpha_{n+1} = \alpha_n$, so that the sequence $\alpha_1, \alpha_2, \alpha_3, \cdots$ determines an element α^* of the inverse limit group $\bar{H}_k(M;S,T)$. We define the linear map h by putting $h\alpha = \alpha^*$. To prove Theorem 7*u* we have only to show that h is an isomorphism of H_k^u onto $\bar{H}_k$. It may be remarked that the fact that h is a mapping onto $\bar{H}_k$ is independent of the triangulability of S. We shall prove that h is a mapping onto first, and then shall prove that $h\alpha = 0$ implies $\alpha = 0$.

Proof that h is onto: Let $\beta_1, \beta_2, \beta_3, \cdots$ be a sequence determining an element β^* in the limit group $\bar{H}_k$, so that $i_n^*\beta_{n+1} = \beta_n$. Let $z_1, z_2, z_3, \cdots$ be a sequence of closed forms such that z_k belongs to the class β_k. We shall show that it is possible to choose z_k' in the same class β_k as z_k in such a way that $z_{k+1}' = z_k'$ near C_k. Once this is done, we have only to put $z = \lim_{k\to\infty} z_k$ and take $\beta \varepsilon H_k^u$ to be the class of z. This done, we have $h\beta = \beta^*$, so that h maps H_k^u onto $\bar{H}_k$.

We make our central construction inductively. Put $z_1' = z_1$. Suppose that z_k' is defined for $k < n$ in such a way as to ensure (a) $z_k' \varepsilon \beta_k$, $k < n$ and (b) $z_k' = z_{k-1}'$ in a neighborhood of C_{k-1}, $k < n$. Then since $i_{n-1}^* z_n$ is also in the class β_{n-1}, there exists forms a and b such that $z_n = z_{n-1}' + da + b$, where $a \equiv 0$ in a neighborhood of T_{n-1}^u and $b \equiv 0$ in a neighborhood of S_{n-1}^u. Define a' and b' in such a way that a' and b' are C^∞ forms on M and so that (i) $a' = 0$ in a neighborhood of C_{n-1}; (ii) $a' = a$ in a neighborhood of T_n^u; (j) $b' = 0$ in a neighborhood of C_{n-1}; (jj) $b' = b$ in a neighborhood of S_n^u. This is possible since $a \equiv 0$ in a neighborhood of $C_{n-1} \cap T_{n-1}^u$, and $b \equiv 0$ in a neighborhood of $C_{n-1} \cap S_n^u = S_{n-1}^u$. Now put

$$z_n' = z_{n-1}' + da' + b'.$$

Then $z_n' = z_{n-1}'$ in a neighborhood of C_{n-1}, and also

$$z_n' = z_n + d(a' - a) + (b' - b).$$

Since $a' - a \equiv 0$ in a neighborhood of T_n^u and $b' - b \equiv 0$ in a neighborhood of S_n^u, z_n' and z_n both belong to the class β_n. Thus we have completed the inductive step in the construction of the sequence $z_1', z_2', z_3', \cdots$. Consequently, we have shown that h maps H_k^u onto $\bar{H}_k$.

Proof that $h\alpha = 0$ implies $\alpha = 0$. At this point, we must introduce the hypothesis that S is a subcomplex of a triangulation of M. Since the particular sequence C_n of compact sets that we use to exhaust M does not affect the limit-group (cf. Eilenberg-Steenrod [2], pp. 219-220) we can assume, without loss of generality, that C_n is a finite subcomplex of M. Let z be an element of the class α. To say that $h\alpha = 0$ is to say that z belongs to the class zero in each of the groups $H_k(M; S_n^u, T_n^u)$. Thus, for each n, there exists a form a_n which vanishes in a neighborhood of T_n^u such that $z \equiv da_n$ in a neighborhood of S_n^u. Now, $d(a_{n+1} - a_n) = 0$ in a neighborhood of S_n^u, so that $a_{n+1} - a_n$ belongs to a certain cohomology class; i. e., $a_{n+1} - a_n$ determines an element of $H_k(M; S_n^u, T_n^u)$. Our first aim is to show that we can choose the sequence a_n in such a way that $a_{n+1} - a_n$ determines the zero element of $H_{k-1}(M; S_n^u, T_n^u)$.

Let $A_n^{(1)}$, $n = 1, 2, \cdots$ be the set of all a such that $a \equiv 0$ in a neighborhood of T_n^u and $z \equiv da$ in a neighborhood of S_n^u. Choose $b_1 \varepsilon A_1^{(1)}$. Let $G_n^{(1)} \subseteq H_{k-1}(M; S_1^u, T_1^u)$ be the set of all cohomology classes determined by the elements of $b_1 - A_n^{(1)}$. Then $G_n^{(1)} \supseteq G_{n+1}^{(1)}$, and each G_n is a non-null linear variety. Since S_1^u and T_1^u are finite subcomplexes of M, $H_{k-1}(M; S_1^u, T_1^u)$ is finite dimensional. Hence $\bigcap_{n=2}^{\infty} G_n^{(1)} = G^{(1)}$ is non-null. Let $\alpha_1 \varepsilon G^{(1)}$, and let d_1 be an element of the cohomology class α_1. Put $a_1 = b_1 + d_1$.

Let $A_n^{(2)}$, $n = 2, 3, \cdots$, be the set of all elements $a \varepsilon A_n^{(1)}$ such that $a_1 - a$ belongs to the zero class of $H_k(M; S_1^u, T_1^u)$. It follows from the construction of a_1 that each $A_n^{(2)}$ is non-null. Let $b_2 \varepsilon A_2^{(2)}$, and let $G_n^{(2)} \subseteq H_{k-1}(M; S_2^u, T_2^u)$ be the set of all cohomology classes determined by the elements of $b_2 - A_n^{(2)}$. Then $\bigcap_{n=3}^{\infty} G_n^{(2)} = G^{(2)} \neq \emptyset$. Let $\alpha_2 \varepsilon G^{(2)}$, and let d_2 be an element of the class α_2. Put $a_2 = b_2 + d_2$. Then, continuing inductively, define $a_3, a_4, \cdots$ in such a way that $a_{n+1} - a_n$ always belongs to the zero class of $H_{k-1}(M; S_n^u, T_n^u)$.

Next we will construct a sequence of elements $a_n' \varepsilon A_n^{(1)}$ such that $a_n' = a_{n+1}'$ in a neighborhood of C_n. Once this is done, we have only to take $a = \lim_{n \to \infty} a_n'$ in order to have $a \equiv 0$ in a neighborhood of T and $z \equiv da$ in a neighborhood of S. However, the existence of such an a means precisely that z belongs to the zero class of $H_k^u(M; S, T)$, i. e., that $\alpha = 0$. Thus, all that remains is the construction of the sequence a_n'.

We make the construction inductively. Put $a_1' = a_1$. Suppose that a_m' has been defined for $m < n$ in such a way as to ensure that (a) $a_m' \varepsilon A_m^{(1)}$; (b) $a_m' - a_m$ belongs to the zero class of $H_{k-1}(M; S_m^u, T_m^u)$; (c) $a_m' - a_{m-1}' \equiv 0$ in some neighborhood of C_{m-1}, for $m < n$. Then a_n can be written in the form $a_n = a_{n-1}' + db + c$, where $b \equiv 0$ in a neighborhood of T_{n-1}^u, and $c \equiv 0$ in a neighborhood of S_{n-1}^u. Let b' and c' be C^∞ forms on M such that (i) $b' = 0$ in a neighborhood of C_{n-1}; (ii) $b' = b$ in a neighborhood of T_n^u; (j) $c' = 0$ in a neighborhood of C_{n-1}; (jj) $c' = c$ in a neighborhood of S_n^u. This is possible since $b = 0$ in a neighborhood of $C_{n-1} \cap T_n^u = T_{n-1}^u$ and $c = 0$ in a neighborhood of $C_{n-1} \cap S_n^u = S_{n-1}^u$. Put $a_n' = a_{n-1}' + db' + c'$. Then $a_n' = a_n + d(b - b') + (c - c')$. It follows readily that $a_n' \varepsilon A_n^{(1)}$, that $a_n' - a_n$ is in the zero class of $H_{k-1}(M; S_n^u, T_n^u)$, and that $a_n' - a_{n-1}' \equiv 0$ in a neighborhood of C_{n-1}. This concludes our inductive step, and with it, our proof of Theorem 7*u*.

Now if M is triangulated, and S and T are subcomplexes of M, then it is known that the singular homology group $\mathfrak{H}_c^n(S, T)$ of finite cycles arises as the limit group of the direct sequence of homology groups obtained from the sequence of identity maps i_n of [*]. Moreover, it is known that the singular homology group $\mathfrak{H}_u^n(S, T)$ of locally finite cycles arises as the limit group of the inverse sequence of homology groups obtained from the sequence of identity maps j_n of [**]. Now, in the case in which all terms are finite dimensional vector spaces, the inverse limit of an inverse sequence is dually paired to the direct limit of the direct sequence of dual groups. (Cf. Lefschetz [3], 120. 81, p. 67, and also 6, pp. 74-83, especially 33, p. 83). We have therefore

THEOREM 8c. *Let M be a triangulable C^∞ manifold, and let (S, T) be a pair of subcomplexes of M. Then $H_k^u(M; S, T)$ and $\mathfrak{H}_c^k(S, T)$ are dually paired under the natural pairing defined by the integral of a k-form over a finite k-chain.*

THEOREM 8u. *Let M be a triangulable C^∞ manifold, and let (S, T) be a pair of subcomplexes of M. Then $H_k^c(M; S, T)$ and $\mathfrak{H}_u^k(S, T)$ are dually paired under the natural pairing defined by the integral of a k-form with compact carrier over a locally finite k-chain.*

YALE UNIVERSITY.

BIBLIOGRAPHY.

[1] S. Eilenberg, "Singular homology in differentiable manifolds," *Annals of Mathematics*, vol. 48 (1947), p. 670.

[2] S. Eilenberg and N. Steenrod, *Foundations of Algebraic Topology*, Princeton, 1952.

[3] S. Lefshetz, *Algebaic Topology*, American Mathematical Society Colloquium Publications, vol. 27 (1942).

ASYMPTOTIC INTEGRATIONS OF LINEAR DIFFERENTIAL EQUATIONS.*

By PHILIP HARTMAN and AUREL WINTNER.

Introduction.

In the linear system of differential equations

$$\text{(i)} \qquad y' = (J + G(t))y,$$

where y' denotes dy/dt and y is a (column) vector with d components, let J be a d by d matrix of complex numbers, and $G(t)$ a d by d matrix of continuous (possibly complex-valued) functions $g_{jk}(t)$ on the half-line $0 \leqq t < \infty$. Let $\lambda_1, \cdots, \lambda_d$ denote the characteristic numbers of J and $\mu_j = \operatorname{Re} \lambda_j$ the real part of λ_j. Finally, put

$$\text{(ii)} \qquad | G(t) | = \sum_{j=1}^{d} \sum_{k=1}^{d} | g_{jk} |.$$

Part I of this paper deals with the asymptotic integration of (i) on a logarithmic scale. If

$$\text{(iii)} \qquad | G(t) | \to 0 \text{ as } t \to \infty,$$

then (i) has d linearly independent solutions $y(t) = y_1(t), \cdots, y_d(t)$ with the property that, if $| y |$ denotes the length of the vector y,

$$\text{(iv)} \qquad \lim_{t \to \infty} t^{-1} \log | y(t) | = \mu,$$

where $\mu = \mu_m$, $y(t) = y_m(t)$ and $m = 1, \cdots, d$. This was proved by Perron (cf. [8], pp. 158-159) with the lim replaced by lim sup, but his method can be used to supply (iv) as it stands, cf. [7].

In view of known results ([2], [3]) for the case when (i) is equivalent to a second order differential equation, it is natural to ask under what conditions, weaker than (iii), will the assertion concerning (iv) remain true. In this direction, it will be shown that (iii) can be relaxed to

$$\text{(v)} \qquad \underset{0 \leqq V < \infty}{\text{l. u. b.}} \int_{U}^{U+V} | G(t) | \, dt/(1 + V) \to 0 \text{ as } U \to \infty.$$

---

* Received October 21, 1954.

It is clear that (iii) implies (v) and also that (v) holds if $|G(t)|$ is of class $L^p(0,\infty)$ for some $p \geqq 1$. (When $\mu_1 = \cdots = \mu_d$, condition (v) can be relaxed to

$$T^{-1} \int_0^T |G(t)|\, dt \rightarrow 0,$$

as $T \rightarrow \infty$.)

The fact that (v) implies the statement concerning (iv) is a corollary of (*) in Section 2 below. In contrast to Perron's proof, the proof of (*) will not depend on successive approximations. This leads, among other things, to an extension of (*) to non-linear systems under conditions less severe than those used by Perron in [8] (this extension will be carried out in a forthcoming paper). It also leads to theorems on the asymptotic integration of (i) under conditions different from those in standard theorems; cf. Part III.

Parts II and III will deal with asymptotic formulae for solutions of (i), rather than for the logarithms of their magnitudes. In Part II, there will be deduced (Theorem (**)) a refinement of a theorem obtained by Dunkel in his dissertation [1], written under the direction of Bôcher. Particular cases of this theorem have been rediscovered repeatedly (possibly because the result was published in a periodical which is not readily accessible). On the other hand, as far as we know, the only proof of Dunkel's complete theorem in the literature is that of Dunkel himself. His proof is based on simple ideas but is rather difficult to read because of complicated computations involved in his successive approximations. It seems to be worthwhile to give a proof which avoids successive approximations. After a suitable change of variables, the main theorem (**) of Part II will be obtained as a corollary of the results of Part I.

If the independent variable x in Dunkel's theorem is replaced by $t = -\log x$, his assertion implies that if $y_0(t) = c \exp \lambda t$, where $c \neq 0$ is a constant vector, is a solution of the system

(vi) $$y' = Jy,$$

if the elementary divisors of J belonging to the characteristic numbers λ_j satisfying $\mu_j = \mu$ $(= \operatorname{Re} \lambda)$ are simple, and if

(vii) $$\int^{\infty} |G(t)|\, dt < \infty,$$

then (i) has a solution $y = y(t)$ satisfying, as $t \rightarrow \infty$,

(viii) $$y(t) - y_0(t) = o(e^{\mu t}).$$

More generally, if $y_0 = (c_1 t^k + c_2^{k-1} + \cdots + c_{k+1}) \exp \lambda_m t$, where $c_1, \cdots, c_{k+1}$ are constant vectors and $c_1 \neq 0$, is a solution of (vi), if h_* denotes the maximum multiplicity of the elementary divisors of J belonging to the eigenvalues λ_j satisfying $\mu_j = \mu_m$, and if

$$\int^{\infty} t^{h_*-1} \, | G(t) | \, dt < \infty, \tag{ix}$$

then (i) has a solution $y = y(t)$ satisfying, as $t \to \infty$,

$$y(t) - y_0(t) = o(t^{k-1} e^{\mu t}). \tag{x}$$

It will be shown below that if $h_0 \geqq h_*$ and if condition (ix) is strengthened to

$$\int^{\infty} t^{h_0-1} \, | G(t) | \, dt < \infty, \qquad (h_0 \geqq h_*), \tag{xi}$$

then assertion (x) can be improved to

$$y(t) - y_0(t) = o(t^{h_0 - h_* + k - 1} e^{\mu t}). \tag{xii}$$

In particular, if $h_0 \geqq 2h_* - 1$, then (viii) holds. For example, if

$$\int^{\infty} t^{2(d-1)} \, | G(t) | \, dt < \infty, \tag{xiii}$$

then to every solution $y = y_0(t)$ of (vi) satisfying (iv), there exists a solution $y = y(t)$ of (i) satisfying (viii).

The assertion (**) in Section 11 refining Dunkel's theorem is stated for the case when J is in its Jordan normal form. This permits a refinement of the assumption (xi) and the assertion (xii). For example, if the d-th order, linear differential equation for the scalar x,

$$x^{(d)} + f_1(t) x^{(d-1)} + \cdots + f_{d-1}(t) x' + f_d(t) x = 0, \tag{xiv}$$

is written as a system of first order linear differential equations for the vector $y = (x, x', \cdots, x^{(d-1)})$, then the following fact will follow from (**): If the (possibly complex-valued) coefficients $f_1(t), \cdots, f_d(t)$ of (xiv) are continuous for large t and satisfy

$$\int^{\infty} t^{j-1} \, | f_j(t) | \, dt < \infty \qquad (j = 1, \cdots, d) \tag{xv}$$

and if $0 \leqq k \leqq d - 1$, then (xiv) has a solution $x = x(t) = x_k(t)$ satisfying, as $t \to \infty$,

$$x \sim t^k, \; x' \sim k t^{k-1}, \cdots, x^{(k)} \sim k! \tag{xvi}$$

and

$$\text{(xvii)} \qquad x^{(k+1)} = o(t^{-1}), \cdots, x^{(d-1)} = o(t^{k-d+1}).$$

On the other hand, by the statement of Dunkel's theorem (cf. [1], pp. 368-370), the condition (xv) must be replaced by the stronger requirement that $t^{d-1} \mid f_j(t) \mid$ is of class L for $j = 1, \cdots, d$ and the assertion in (xvii) by the weaker statement that

$$\text{(xviii)} \qquad x^{(k+1)} = o(1), \cdots, x^{(d-1)} = o(1)$$

as $t \to \infty$. (Of course, when the coefficients $f_j(t)$ are not "close to 0" in the sense of (xv), but are "near" to other constants, then (**) still furnishes an asymptotic integration of (xiv).)

In Section 17 at the end of Part II, Theorem (**) will be used to obtain the following theorem dealing with a system of the form

$$\text{(xix)} \qquad y' = (P(t) + G(t))y,$$

in which the constant matrix J in (i) is replaced by a continuous periodic matrix $P(t)$, say of period 1, so that $P(t+1) = P(t)$: If $y = y_0(t)$ is a solution of

$$\text{(xx)} \qquad y' = P(t)y$$

satisfying (iv) and if $G(t)$ is continuous for large t and satisfies (xiii), then (xix) has a solution $y = y(t)$ satisfying (viii) as $t \to \infty$. (In particular, if $y_0 = e^{\lambda t}p(t)$ is a solution of (xx), where $p(t) \not\equiv 0$ is a vector of period 1, then (xiii) implies that (xix) has a solution $y = y(t)$ satisfying (viii), where $\mu = \operatorname{Re} \lambda$.)

Part III will deal with analogues of the result of Part II, when the conditions (vii), (ix) or (x) are replaced, respectively, by

$$\int^{\infty} \mid G(t) \mid^{\gamma} dt < \infty, \quad \int^{\infty} t^{\gamma h_{*}-1} \mid G(t) \mid dt < \infty \text{ or } \int^{\infty} t^{\gamma h_0 - 1} \mid G(t) \mid dt < \infty,$$

where γ is some number on the range $1 \leqq \gamma \leqq 2$. The proof of the main results of Part III depend on a modification of the proof of (*) in Part I.

The results obtained generalize the theorem (VII) in [2], p. 575, dealing with a second order linear differential equation

$$\text{(xxi)} \qquad x'' - f(t)x = 0,$$

where $f(t)$ is real-valued and "nearly" 1 for large t. The theorems to be obtained for the system (i) allow an extension of theorem (VII) of [2] to cases in which $f(t)$ is complex-valued and "near" a complex number λ, which is not 0 or a negative real number, and to cases in which $|f(t)| \to \infty$ as $t \to \infty$. For example, it is shown that if λ is a complex number which is neither negative nor zero (so that $\operatorname{Re} \lambda^{\frac{1}{2}} \neq 0$) and if $f(t)$ is a complex-valued continuous function for large t satisfying

$$\int^{\infty} |f(t) - \lambda|^{\gamma} \, dt < \infty \text{ for some } \gamma, \qquad 1 \leqq \gamma \leqq 2,$$

then (xxi) has a pair of solutions satisfying, as $t \to \infty$,

$$x \sim \exp \pm \int^{t} \tfrac{1}{2}(\lambda^{\frac{1}{2}} + \lambda^{-\frac{1}{2}} f(u))\, du, \qquad x' \sim \pm \lambda^{\frac{1}{2}} x.$$

It is also shown that if $f(t)$ is a complex-valued continuously differentiable function such that

$$\text{(xxii)} \qquad \operatorname{Re} f^{\frac{1}{2}}(t) \neq 0, \qquad \int^{\infty} |\operatorname{Re} f^{\frac{1}{2}}(t)| \, dt = \infty, \qquad f' = o(|f| \, |\operatorname{Re} f^{\frac{1}{2}}|),$$

then (xxi) has a pair of solutions satisfying, as $t \to \infty$,

$$\text{(xxiii)} \qquad x' \sim \pm f^{\frac{1}{2}} x.$$

If the last condition in (xxii) is replaced by

$$\text{(xxiv)} \qquad \int^{\infty} |f'/f|^{\gamma} \, |\operatorname{Re} f^{\frac{1}{2}}|^{\gamma - 1} \, dt < \infty \text{ for some } \gamma, \text{ where } 1 \leqq \gamma \leqq 2,$$

then (xxi) has solutions satisfying (xxiii) and

$$\text{(xxv)} \qquad x \sim f^{-\frac{1}{4}}(t) \exp \pm \int^{t} f^{\frac{1}{2}}(u)\, du,$$

where $f^{-\frac{1}{4}} = (f^{\frac{1}{2}})^{-\frac{1}{2}}$. Somewhat different asymptotic formulae follow if the three conditions of (xxii) are retained and (xxiv) is replaced by

$$\text{(xxvi)} \qquad \int^{\infty} |d(f'/f^{\frac{3}{2}})| < \infty.$$

In this case, (xxi) has a pair of solutions satisfying (xxiii) and

$$\text{(xxvii)} \qquad x \sim f^{-\frac{1}{4}}(t) \exp \pm \int^{t} f^{\frac{1}{2}} (1 + f'^{2}/16 f^{3})^{\frac{1}{2}} \, du.$$

While the assertion concerning (xxiv) has no analogue in the elliptic case (say, when $f(t)$ is negative), the assertion concerning (xxvi) does: Let $f(t)$ be a continuously differentiable function for large t satisfying (xxviii) $\qquad f(t) \neq 0$ and $\exp \pm i \int^{t} f^{\frac{1}{2}}(1 + f'^2/16f^3)^{\frac{1}{2}}\, du = O(1)$,

(xxvi) and

(xxix) $$\lim_{t\to\infty} (1 + f'^2/16f^3) \neq 0.$$

Then the differential equation

(xxi bis) $$x'' + f(t)x = 0$$

has a pair of solutions satisfying, as $t \to \infty$,

$$x \sim f^{-\frac{1}{4}}(t) \exp \pm i \int^{t} f^{\frac{1}{2}}(1 + f'^2/16f^3)^{\frac{1}{2}}\, du, \qquad x' \sim \text{const.}\, f^{\frac{1}{2}}x,$$

where the const. depends on the value of the limit (xxix).

For corresponding results under different conditions on f, cf. [10].

Part I. The logarithmic scale.

1. Notation. Without loss of generality, it can be supposed that the constant matrix J is in its Jordan normal form; so that J consists of a number of "blocks," say g "blocks," $J(1), \cdots, J(g)$, on its diagonal, where $J(j)$ is an $h(j)$ by $h(j)$ matrix with a certain number $\lambda = \lambda(j)$ on its diagonal and, if $h(j) > 1$, the number 1 on its subdiagonal, where $h(1) + \cdots + h(g) = d$; the numbers $\lambda(1), \cdots, \lambda(g)$ need not be distinct.

The system (vi) reduces to

(1_{j1}) $$y^{j1\prime} = \lambda y^{j1},$$

(1_{jk}) $$y^{jk\prime} = \lambda y^{jk} + y^{j\,k-1}, \qquad k = 2, \cdots, h(j),$$

where $\lambda = \lambda(j)$; $j = 1, \cdots, g$; $h(1) + \cdots + h(g) = d$; $y^m = y^{jk}$ if $m = h(1) + \cdots + h(j-1) + k$ for $k = 1, \cdots, h(j)$. The equations (1_{jk}), $k > 1$, are missing if $h(j) = 1$.

Similarly, (i) can be written as

(2_{j1}) $$y^{j1\prime} = \lambda y^{j1} + g_{j1\,\alpha\beta}\, y^{\alpha\beta},$$

(2_{jk}) $$y^{jk\prime} = \lambda y^{jk} + y^{j\,k-1} + g_{jk\,\alpha\beta}\, y^{\alpha\beta}, \qquad k = 2, \cdots, h,$$

where $\lambda = \lambda(j)$, $h = h(j)$; $j = 1, \cdots, g$; the element $g_{m\gamma}(t)$ of the matrix $G(t)$ is written as $g_{m\gamma} = g_{jk\,\alpha\beta}$ if $m = h(1) + \cdots + h(j-1) + k$, $\gamma = h(1) + \cdots + h(\alpha - 1) + \beta$ for $k = 1, \cdots, h(j)$ and $\beta = 1, \cdots, h(\alpha)$;

finally, the summation convention is used for the indices α, β; so that the last term of (2_{jk}) means

$$g_{jk\,\alpha\beta}y^{\alpha\beta} \equiv \sum_{\alpha=1}^{g}\sum_{\beta=1}^{h(\alpha)} g_{jk\,\alpha\beta}y^{\alpha\beta}.$$

It can be supposed that the enumeration has been chosen so that $\mu(j) = \operatorname{Re}\lambda(j)$ satisfies

$$\mu(1) \leqq \cdots \leqq \mu(g). \tag{3}$$

Let the distinct numbers in the set (3) be denoted by $\mu^1 < \mu^2 < \cdots < \mu^f$. For a given integer m, where $1 \leqq m \leqq f$, an integer j on the range $1 \leqq j \leqq g$ will be denoted by p, q or r according as $\mu(j) < \mu^m$, $\mu(j) = \mu^m$ or $\mu(j) > \mu^m$:

$$\mu(j) \lesseqqgtr \mu^m \text{ according as } j = p,\ q \text{ or } r. \tag{4}$$

In terms of these notations, put

$$L_m = \sum_q \sum_{k=1}^{h(q)} |\, y^{qk}\,|^2; \tag{5}$$

so that L_m is the sum of the squares of the absolute values of the components y^j of y corresponding to which $\operatorname{Re}\lambda_j = \mu^m$. Thus $|\,y\,|^2 = L_1 + \cdots + L_f$.

2. The main theorem. The main result of Part I can be stated as follows:

(*) *Let $G(t)$ be continuous on $0 \leqq t < \infty$ and let*

$$\operatorname*{l.\,u.\,b.}_{0 \leqq V < \infty} \int_U^{U+V} |\, G(t)\,|\, dt/(1+V) \to 0 \text{ as } U \to \infty, \tag{6}$$

that is, for every set of indices $(jk\,\alpha\beta)$, let

$$\operatorname*{l.\,u.\,b.}_{0 \leqq V < \infty} \int_U^{U+V} |\, g_{jk\,\alpha\beta}\,|\, dt/(1+V) \to 0 \text{ as } U \to \infty.$$

Let $1 \leqq m \leqq f$. Then there exists a positive number T with the property that to any number $t_0 \geqq T$ and to any set of $\sum h(q)$ numbers $y_0{}^{qk}$, not all 0, there belongs a unique set of $\sum h(r)$ numbers $y_0{}^{rk}$ such that the solution $y = y(t)$ of (2), determined by the initial conditions

(7_1) $y^{pk}(t_0) = 0$; (7_2) $y^{qk}(t_0) = y_0{}^{qk}$; (7_3) $y^{rk}(t_0) = y_0{}^{rk}$,

satisfies, as $t \to \infty$,

$$L_j(t) = o(L_m(t)) \text{ if } j \neq m \tag{8}$$

and

$$\log |y(t)| = (\mu^m + o(1))t. \tag{9}$$

It will follow from the proof of (*) that if $0 \leqq \eta < 1$ and if $T = T_\eta$ is sufficiently large, then (7_1) can be replaced by $y^{pk}(t_0) = y_0^{pk}$, where y_0^{pk} are arbitrary numbers satisfying $\Sigma \mid y_0^{pk} \mid \leqq \eta \Sigma \mid y_0^{qk} \mid$.

It is understood that the subset of initial conditions (7_1) or (7_3) is vacuous if $m = 1$ or $m = f$. Both of these sets of initial conditions are vacuous in the case

$$\mu_1 = \cdot \cdot \cdot = \mu_d, \tag{10}$$

where $f = 1$.

(I) *In the case* (10), *condition* (6) *in* (*) *can be relaxed to*

$$T^{-1} \int^T | G(t) | \, dt \rightarrow 0 \text{ as } T \rightarrow \infty, \tag{11}$$

that is, to

$$T^{-1} \int^T | g_{jk\,\alpha\beta} | \, dt \rightarrow 0 \text{ as } T \rightarrow \infty$$

(*for every set of indices* $(jk\,\alpha\beta)$).

Since (*) shows that there are $\Sigma h(q)$ linearly independent solutions of (i) satisfying (9), it is seen that (*) implies the theorem of Perron mentioned in the Introduction.

Assertion (*) contains as a corollary the following theorem of Perron [8], p. 765:

(II) *Let* J *be a diagonal matrix with diagonal elements* $\lambda_1, \cdot \cdot \cdot, \lambda_d$ *satisfying* $\mu_1 < \cdot \cdot \cdot < \mu_d$, *where* $\mu_j = \operatorname{Re} \lambda_j$. *Let* $G(t)$ *be a continuous matrix on* $0 \leqq t < \infty$ *satisfying*

$$G(t) \rightarrow 0 \text{ as } t \rightarrow \infty. \tag{12}$$

Then for any m, *where* $1 \leqq m \leqq d$, *the system* (i) *has a solution* $y = y(t) = (y^1(t), \cdot \cdot \cdot, y^d(t))$, *the components* $y^j(t)$ *of which satisfy, as* $t \rightarrow \infty$,

$$y^j(t) = o(y^m(t)) \text{ if } j \neq m, \tag{13}$$

$$y^{m\prime}(t)/y^m(t) \rightarrow \lambda_m. \tag{14}$$

In order to deduce this from (*), note that (12) implies (6); so that, under the conditions of (II), assertion (*) is applicable to (i). Hence (i) has a solution satisfying (8), (9). Since J is a diagonal matrix and $\mu_1 < \cdot \cdot \cdot < \mu_d$, the sum in (5) reduces to a single term. Hence, the relations

(8) and (13) are equivalent. The fact that J is a diagonal matrix means that the differential equation for y^m in (i) is

$$y^{m\prime} = \lambda_m y^m + \sum_{j=1}^{d} g_{mj} y^j. \tag{15}$$

Hence (14) follows from (12) and (13).

In (II), in order to assure the existence of a solution satisfying (13), (14) for a given m, the condition that J is a diagonal matrix can be relaxed to the assumption that $\mu_j \neq \mu_m$ for all $j \neq m$ (so that there is only one elementary divisor belonging to λ_k with $\operatorname{Re}\lambda_k = \mu_m$ and this elementary divisor is simple).

An easy example shows that (II) has no simple analogue if $\mu_1 = \cdots = \mu_d$. In fact, if $d = 2$, $\lambda_1 = \lambda_2 = 0$ and $g_{12} = g_{22} = 0$, then (2) becomes

$$y^{1\prime} = g_{11} y^1, \qquad y^{2\prime} = g_{21} y^1.$$

The solutions of this system are

$$y^1 = c_1 \exp \int^t g_{11} ds, \qquad y^2 = c_1 \int^t g_{21} (\exp \int^s g_{11} du) ds + c_2.$$

It is clear that g_{11} and g_{21} can be chosen so as to satisfy $g_{jk} \to 0$ as $t \to \infty$ but to leave quite arbitrary the asymptotic behavior of the ratio y^2/y^1 (for a solution in which $c_1 \neq 0$).

3. **Riccati equations.**

The proof of (*) will be similar to the proof of a particular case of it given in [2]. It will depend on the use of Riccati differential equations as "majorants." (For applications of Riccati equations as "majorants" for certain non-linear differential equations, cf. [5], pp. 353-355.)

LEMMA 1. *Let $f(t)$ be a continuous real-valued function on $0 \leqq t < \infty$ satisfying*

$$\underset{0<S<\infty}{\text{l. u. b.}} \left| \int_T^{T+S} f(t)dt/(1+S) \right| \to 0 \text{ as } T \to \infty, \tag{16}$$

and let $c \neq 0$ be a real number. Then, if $r = r(t)$ is a real-valued solution of

$$r' = cr(1-r) + f(t) \tag{17}$$

for large t,

$$r(\infty) = \lim_{t\to\infty} r(t) \text{ exists and is 0 or 1.} \tag{18}$$

Furthermore, (17) *has, for large* t, *solutions* $r(t)$ *satisfying* $r(\infty)=0$ *and solutions* $r(t)$ *satisfying* $r(\infty)=1$.

This theorem is contained in Theorems (I), p. 560 and (II), p. 564 and in their applications on pp. 570-571 in [2]. The proofs given there imply the following fact:

LEMMA 2. *Let* $c>0$ *and let* $f(t)$ *satisfy the assumption of Lemma* 1. *Then there belongs to every* ϵ, $0<\epsilon<1$, *a* $T_\epsilon=T(\epsilon,f)$ *with the properties that if* $r=r(t)$ *is solution of* (17) *on some* t*-interval* $[t_0,t^0]$, $t_0\geqq T_\epsilon$, *and satisfies*

$$r(t)\geqq 2\epsilon \tag{19}$$

for $t=t_0$, *then* $r(t)$ *can be continued over* $t_0\leqq t<\infty$ *and satisfies*

$$r(t)>\epsilon \tag{20}$$

for all $t\geqq t_0$ (*in particular,* $r(\infty)=1$).

These lemmas have the following analogues in which the differential equation (17) is replaced by a differential inequality:

LEMMA 1 bis. *Let* c *and* $f(t)$ *satisfy the conditions of Lemma* 2 *and, for large* t, *let* $v=v(t)$ *be a continuously differentiable function satisfying*

$$v'\geqq cv(1-v)+f(t). \tag{21}$$

Then either

$$\infty\geqq\liminf_{t\to\infty}v(t)\geqq 1 \tag{22}$$

or

$$-\infty\leqq\limsup_{t\to\infty}v(t)\leqq 0. \tag{23}$$

The analogue of Lemma 2 is

LEMMA 2 bis. *Let* c, $f(t)$ *and* $v(t)$ *satisfy the assumptions of Lemma* 1 bis *and let* ϵ, T_ϵ *be the same as in Lemma* 2. *If, in addition,*

$$v(t)\geqq 2\epsilon \tag{24}$$

for some $t=t_0\geqq T_\epsilon$, *then*

$$v(t)>\epsilon \tag{25}$$

for all $t\geqq t_0$ (*and, therefore,* (22) *holds*).

Lemmas 1 bis and 2 bis follow from Lemma 2 and the fact that if $r(t)$ is a solution of (17) on some t-interval $[t_0,t^0]$ and if $v(t)\geqq r(t)$ at

$t = t_0$, then $v(t) \geqq r(t)$ for $t_0 \leqq t \leqq t^0$. Thus, if (23) fails to hold, then $v(t) \geqq 2\epsilon > 0$ for arbitrarily large $t = t_0$. Hence, if $r(t)$ is determined by the initial condition $r(t_0) = 2\epsilon$, then $r(t)$ can be defined for $t_0 \leqq t < \infty$, (20) is satisfied and $r(\infty) = 1$, while $v(t) \geqq r(t)$, and so (22) holds.

4. Proof of (*). Preliminaries. In the proof of (*), it can be supposed that $f > 1$. For the proof in this case will show that the case $f = 1$ (that is, the case (10) in (I)) is quite simple; cf. the deduction of (41) from (40) below.

Let a be a fixed number on the range $0 < a < 1$, so small that

$$(26) \qquad c = c_m = 2(\mu^m - \mu^{m-1} - 2a) > 0 \text{ for } m = 2, \cdots, f.$$

Let $|G(t)|$ denote the scalar function defined by

$$(27) \qquad |G(t)| = \sum_{j=1}^{d} \sum_{k=1}^{d} |g_{jk}(t)| \equiv \sum_{j=1}^{g} \sum_{k=1}^{h(j)} \sum_{\alpha=1}^{g} \sum_{\beta=1}^{h(\alpha)} |g_{jk\,\alpha\beta}(t)|;$$

so that (6) holds as a scalar relation if $|G(t)|$ is interpreted to be the function (27).

According as $k = 1$ or $k > 1$, (2_{jk}) shows that, for any solution $y = y(t)$ of (2), the absolute value of

$$(28) \qquad (|y^{jk}|^2)' - 2\mu(j)|y^{jk}|^2$$

is majorized by $|G(t)| \cdot |y(t)|^2$ or by $2|y^{jk}y^{j\,k-1}| + |G(t)| \cdot |y|^2$. Since $2|bc| \leqq ab^2 + a^{-1}c^2$ for any pair of numbers b and c, it follows that $2|y^{jk}y^{j\,k-1}| \leqq a|y^{jk}|^2 + a^{-1}|y^{j\,k-1}|^2$. Hence, if the resulting inequality for the modulus of (28) is multiplied by a^{2k} and added for $j = q$, $k = 1, \cdots, h(q)$, it is seen that

$$(29) \qquad |L_{ma}' - 2\mu^m L_{ma}| \leqq 2aL_{ma} + da^{-2d}|G(t)|(L_{1a} + \cdots + L_{fa}),$$

where

$$(30) \qquad L_{ma} = \sum_{q} \sum_{k=1}^{h(q)} a^{2k}|y^{qk}|^2.$$

Thus (5) and $0 < a < 1$ imply that

$$(31) \qquad L_{ma} \leqq L_m \leqq a^{-2d} L_{ma}.$$

For $1 \leqq m < f$, put

$$(32) \qquad M = M_m = \sum_{j=1}^{m-1} L_{ja} \text{ and } N = N_m = \sum_{j=m}^{f} L_{ja}.$$

Then (29) shows that

$$(33_m) \qquad M' \leqq 2(\mu^{m-1} + a)M + d^2a^{-2d}|G(t)|(M + N),$$

$$(34_m) \qquad N' \geqq 2(\mu^m - a)N - d^2a^{-2d}|G(t)|(M + N).$$

If the solution $y = y(t)$ under consideration is not identically 0, so that $M + N \geqq a^{2d} | y |^2 > 0$ for all t, define $v = v_m(t)$ by

$$v = N/(N + M). \tag{35}$$

Clearly,

$$0 \leqq v \leqq 1. \tag{36}$$

The relations $v' = (MN' - M'N)/(M + N)^2$ and $v(1 - v) = MN/(M + N)^2$ and the inequalities (33), (34) give

$$v' \geqq c_m v(1 - v) - d^2 a^{-2d} | G(t) |, \tag{37}$$

where c_m is defined as in (26).

In view of (12) and (26), the function $f(t) = - d^2 a^{-2d} | G(t) |$ and the constant $c = c_m$ in the differential inequality (37) satisfy the conditions of Lemmas 1 bis and 2 bis in Section 3.

5. The case $m = f$. The proof for the case $m = f$ can be completed at once. For, in this case, the set of conditions (7_3) belonging to y^{rk} is vacuous. Let $y = y(t)$ be the solution of (2) determined by (7_1)-(7_2). Then $y(t) \not\equiv 0$, since not all of the numbers $y_0{}^{qk}$ are 0. Hence $v(t)$ is defined and is a continuously differentiable function on $0 \leqq t < \infty$ satisfying the differential inequality (37). Since (7) implies $M(t_0) = 0$ and $N(t_0) > 0$, it follows that $v(t_0) = 1$, by (35). Thus Lemma 1 bis shows that if t_0 is sufficiently large, then $\liminf_{t \to \infty} v(t) = 1$. From (36),

$$v(t) \to 1 \text{ as } t \to \infty. \tag{38}$$

The definitions (32), (35) and $m = f$ show, therefore, that, as $t \to \infty$,

$$L_{ja}(t) = o(L_{ma}(t)) \text{ if } j \neq m. \tag{39}$$

In view of (31), this proves assertion (8). By (29) and (39),

$$| L_{ma}' - 2\mu^m L_{ma} | \leqq \{2a + da^{-2d}(1 + o(1)) | G(t) |\} L_{ma}, \tag{40}$$

since $m = f$. Thus (11) implies that, as $t \to \infty$,

$$| t^{-1} \log L_{ma} - 2\mu^m | \leqq 2a + da^{-2d} o(1)$$

or, by (31),

$$| t^{-1} \log L_m - 2\mu^m | \leqq 2a + da^{-2d} o(1) - t^{-1} \log a^{2d}.$$

Consequently,

$$\limsup_{t \to \infty} | t^{-1} \log L_m - 2\mu^m | \leqq 2a. \tag{41}$$

Since $a > 0$ can be chosen arbitrarily small, (9) follows from (8) and (41). This proves the case $m = f$ of (*).

6. The case $m < f$. If $m < f$, introduce, in addition to (35), the auxiliary function $w = v_{m+1}(t)$ defined by

$$w = N_{m+1}/(N_{m+1} + M_{m+1}). \tag{42}$$

Then (42) satisfies

$$0 \leqq w \leqq 1 \tag{43}$$

and the case of (37) in which m is replaced by $m + 1$, that is,

$$w' \geqq c_{m+1} w(1 - w) - d^2 a^{-2d} \mid G(t) \mid. \tag{44}$$

In view of (12) and (26), the function $f(t) = - d^2 a^{-2d} \mid G(t) \mid$ and the constants $c = c_m, c_{m+1}$ in (37) and (44) satisfy the conditions of Lemmas 1 bis and 2 bis in Section 3. Choose ϵ to be $\frac{1}{2}$ and 1 in those assertions and let T be chosen so as to be applicable in the sense of Lemma 2 bis to the differential inequalities (37) and (44). Put $T = \max(T_{\frac{1}{2}}, T_1)$. It will be shown that T has the properties asserted in (*).

The case $\epsilon = 1$ of Lemma 2 bis shows that if $t_0 \geqq T$ and if

$$v(t_0) = 1, \tag{45}$$

then

$$v(t) \geqq \tfrac{1}{2} \tag{46}$$

holds for $t_0 \leqq t < \infty$. Also, the case $\epsilon = \frac{1}{2}$ of Lemma 2 bis shows that if

$$w(t_1) = 0 \text{ for some } t_1 > t_0, \tag{47}$$

then

$$w(t) \leqq \tfrac{1}{2} \tag{48}$$

for $t_0 \leqq t \leqq t_1$, since $w(t) > \frac{1}{2}$ for some t, where $t_0 \leqq t < t_1$, would imply that (47) cannot hold (in fact, that $w(t_1) \geqq \frac{1}{4} > 0$).

7. The solution $y_n(t)$. The existence of the initial conditions (7_3) (belonging to a given $t_0 \geqq T$ and to (7_2)) with the property that (7_1)-(7_3) assure (8) and (9) will be proved by a limit process. If t_0 ($\geqq T$) and if not all of the constants in (7_2) are 0, it will be shown that, corresponding to any $n > t_0$, there exists a (unique) set of $\Sigma\, h(r)$ numbers y_0^{rk} (depending on n) such that the solution $y = y_n(t)$ of (2) determined by (7) satisfies

$$y^{rk}(n) = 0 \text{ for all } r \text{ and for } k = 1, \cdots, h(r). \tag{49}$$

In order to see this, note that, if $y = y(t)$ is a solution of (2) determined by initial conditions of the type

$$y^{jk}(t_0) = 0 \text{ for } j = p, q \tag{50}$$

(and not all $y^{rk}(t_0)$ are 0), then the corresponding function $w(t)$ satisfies $w(t_0) = 1$. Consequently, $w(t) \geqq \frac{1}{2} > 0$ for $t_0 \leqq t < \infty$, in particular for $t = n$. But $w(n) > 0$, that is, $N_{m+1}(n) > 0$, implies that (49) cannot hold. In other words, the $\Sigma_r h(r)$-dimensional manifold of solutions of (2) determined by (50) contains $\Sigma_r h(r)$ solution vectors such that the corresponding $\Sigma_r h(r)$-dimensional vectors with components $y^{rk}(t)$ are linearly independent for $t = n (> t_0)$. Thus, if $y^{rk}(n)$ are any $\Sigma_r h(r)$ numbers, there exists a unique set of $\Sigma_r h(r)$ numbers ${y_0}^{rk}$ with the property that the initial conditions (50) and (7_3) determine a solution such that the values of its components y^{rk} at $t = n$ are the preassigned numbers $y^{rk}(n)$.

Consequently, if $y = y^*(t)$ is the solution of (2) determined by

$$y^{jk}(t_0) = 0 \text{ for } j = p, r \text{ and } y^{qk}(t_0) = {y_0}^{qk},$$

then there is a unique set of initial conditions (7_3) such that if $y = y_*(t)$ is the solution determined by (7_3) and (50), then the components y^{rk} of $y_*(t)$ at $t = n$ assume the same values as the components $-y^{rk}$ of $-y^*(t)$ at $t = n$. Hence, $y = y_*(t) + y^*(t)$ satisfies (7) and (49). This is the solution $y = y_n(t)$, the existence of which was to be proved.

8. Completion of the proof of (*): Existence. Let $v = v^n(t)$ and $w = w^n(t)$, when the latter is defined, denote the auxiliary functions (36), (42) belonging to the solution $y = y_n(t)$. Since $y = y_n(t)$ is determined by initial conditions of the type (7), (7_1) shows that $v = v^n$ satisfies (45), hence (46) for $t_0 \leqq t < \infty$. Since $y = y_n(t)$ satisfies (49), it follows that $w = w^n$ satisfies (47) for $t_1 = n$ $(> t_0)$, hence (48) for $t_0 \leqq t \leqq n$.

The inequality $w^n(t_0) \leqq \frac{1}{2}$ and the fact that $L^n(t_0) = \sum\limits_q \sum\limits_k a^{2k} |\, {y_0}^{qk} \,|^2$ is independent of n show, by (39), that the numbers ${N^n}_{m+1\, a}(t_0)$ are bounded for $n > t_0$. In view of (7), this means that the numbers $|\, y_n(t_0) \,|$ are bounded for $n > t_0$. Consequently, there exists a subsequence of the integers n $(> t_0)$ such that, if n tends to ∞ through this subsequence, then $\lim y_n(t_0) = y_\infty(t_0)$ exists. In view of the locally continuous dependence of solutions of (2) on initial conditions, this implies that

$$\lim y_n(t) = y_\infty(t) \tag{51}$$

exists uniformly on every finite t-interval, as n tends to ∞ through this subsequence of the integers n, and is a solution of (2).

In what follows, n will denote an integer in a subsequence of the integers ($> t_0$) of the type just described.

Clearly, $y = y_\infty(t)$ belongs to initial conditions of the type (7). Hence the corresponding function $v = v^\infty(t)$ satisfies (45) and, therefore, (46) for $t_0 \leqq t < \infty$. It follows from (37), (46) and from Lemma 1 bis that (38) holds for $v = v^\infty(t)$. It is clear from (51) that $w^n(t) \to w^\infty(t)$, as $n \to \infty$, holds uniformly on every finite t-interval. Since $w = w^n(t)$ satisfies (48) for $t_0 \leqq t \leqq t_1$ if $n > t_1$, the function $w = w^\infty(t)$ satisfies (48) for $t_0 \leqq t < \infty$. Consequently (43), (44) and Lemma 1 bis show that

$$w(t) \to 0 \text{ as } t \to \infty \tag{52}$$

holds for $w = w^\infty(t)$.

The definitions of L_{ma}, M_m, N_m, v and w and the relations (38) and (52) imply (39). The assertions (8), (9) of (*) can be deduced from (39), as in the case $m = f$ treated in Section 5 above.

This completes the proof of the existence of a number T and of the numbers $y_0{}^{rk}$ having the properties stated.

9. Completion of the proof of (*): Uniqueness. In order to prove the uniqueness of the set of numbers $y_0{}^{rk}$, suppose that there are two distinct sets. The difference of the two corresponding solutions determined by (7) is a solution $y = y(t) \not\equiv 0$ satisfying (50). The corresponding function $w = w(t)$ satisfies $w(t_0) = 1$, hence $w(t) \to 1$. Thus $M_{m+1}/N_{m+1} \to$ as $t \to \infty$. From (34_{m+1}),

$$\liminf_{t\to\infty} t^{-1} \log N_{m+1}(t) \geqq 2(\mu^{m+1} - a).$$

The argument leading to (41) and then to (9) shows that

$$\liminf_{t\to\infty} t^{-1} \log |\, y(t) \,|^2 \geqq 2\mu^{m+1}.$$

Since this contradicts the fact that $y = y(t)$ is the difference of two solutions satisfying (9), where $\mu^m < \mu^{m+1}$, the proof of (*) is complete.

10. Variants of (*) and (II). In the proof of (*), just completed, the fact that the eignevalues λ_j of J are constants (independent of t) was used only in order to pass from (39) (or, equivalently, from (8)) to (9). Hence the assertion (8) of (*) remains true if the λ_j are functions of t with the property that there exists a constant $a > 0$ such that

$$2(\mu^m - \mu^{m-1} - 2a) \geqq \text{const.} > 0 \quad \text{for} \quad m = 2, \cdots, f. \tag{53}$$

In this case, the number $c = c_m$ of (26) can be replaced by the const. of (53) in the proof of (39) (and/or (8)). In particular, there results the following analogue of (II):

(II*) *Let* $J = J(t)$ *be a diagonal matrix with diagonal elements* $\lambda_1(t), \cdots, \lambda_d(t)$ *satisfying*

$$\mu_j(t) - \mu_{j-1}(t) \geqq \text{const.} > 0 \text{ for } j = 2, \cdots, d, \tag{54}$$

where $\mu_j(t) = \operatorname{Re} \lambda_j(t)$. *Under the assumption* (6) *on* $G(t)$, *the assertion of* (II) *concerning* (13) *holds. If, in addition,* (12) *is satisfied, then* (14) *is valid in the form*

$$y^{m\prime}(t)/y^m(t) \sim \lambda_m(t) \text{ as } t \to \infty. \tag{55}$$

Under the condition (12), rather than (6), this result is contained in a theorem of Perron [8], p. 763.

For a given m, in order to assure the existence of a solution $y = y(t)$ satisfying (12), the condition that $J = J(t)$ is a diagonal matrix satisfying (54) can be relaxed to the assumption that $J = J(t)$ is in its Jordan normal form for every t and that the eigenvalue $\lambda_m(t)$ satisfies

$$| \mu_m(t) - \mu_j(t) | \geqq \text{const.} > 0 \text{ for } j \neq m \tag{56}$$

for large t. Thus, the (single) elementary divisor belonging to $\lambda_m(t)$ is simple but those belonging to $\lambda_j(t)$, where $j \neq m$, need not be.

For applications in Part II, it will be convenient to have a generalization of (II*) dealing with the case when both the diagonal elements λ_j and the 1's on the subdiagonal of J are replaced by functions of t.

Let the system

$$y' = K(t)y \tag{57}$$

be equivalent to

$$y^{j1\prime} = \lambda y^{j1} \tag{58_{j1}}$$

$$y^{jk\prime} = \lambda y^{jk} + \delta y^{j\,k-1}, \qquad k = 2, \cdots, h(j), \tag{58_{jk}}$$

$j = 1, \cdots, g$; $h(1) + \cdots + h(g) = d$; $\lambda = \lambda(j;t)$ and $\delta = \delta(j;t)$. Thus the diagonal elements of $K(t)$ are $\lambda(j;t)$; its subdiagonal elements are $\delta(j;t)$ or 0; and all other elements of $K(t)$ are 0.

Let $\mu(j;t) = \operatorname{Re} \lambda(j;t)$. Let m be an integer on the range $1 \leqq m \leqq g$ and let c denote a positive constant. Relative to m and c, divide the set of integers $j = 1, \cdots, g$ into three sets, members of which will be denoted by p, q, r, respectively, by the requirements that the integer m is a q; that, for some sufficiently small number a, $0 < a < 1$,

$$\mu(p;t) + a \,| \delta p, t) | \leqq \mu(q;t) - a \,| \delta(q,t) | - c \tag{59}$$

for all p, q and large t; and that

$$\mu(r;t) - a\,|\,\delta(r;t)\,| \geqq \mu(q;t) + a\,|\,\delta(q,t)\,| + c \tag{60}$$

for all r, q and large t. Finally, let $L_m(t) = L_m(t;a,c)$ be defined by (5). In terms of this new notation, (*) has the following analogue:

(III) *Let* $G(t)$ *be continuous on* $0 \leqq t < \infty$ *and satisfy* (6). *Let the sets of integers* $\{p\}$, $\{q\}$, $\{r\}$ *and* $L_m(t)$ *be defined as in the preceding paragraph. Then there exists a positive number* T *with the property that to any number* $t_0 \geqq T$ *and to any set of* $\Sigma\, h(q)$ *numbers* $y_0{}^{qk}$, *not all* 0, *there belongs a unique set of* $\Sigma\, h(r)$ *numbers* $y_0{}^{rk}$ *such that the solution of*

$$y' = (K(t) + G(t))y \tag{61}$$

determined by the initial conditions (7) *satisfies, as* $t \to \infty$,

$$y^{jk}(t) = o(L_m^{\frac{1}{2}}(t)) \text{ if } j = p \text{ or } j = r. \tag{62}$$

It is clear that (III) can be extended to the case where $K(t)$ is, for example, a more general triangular matrix. (III) will however suffice for the applications in Part II.

The proof of (III) is very similar to that of (*) and only the beginning of the proof will be indicated. Put

$$L_{ma} = \sum_q \sum_k a^{2k}\,|\,y^{qk}\,|^2; \qquad M = \sum_p \sum_k a^{2k}\,|\,y^{pk}\,|^2; \qquad N = \sum_r \sum_k a^{2k}\,|\,y^{rk}\,|^2.$$

(Here M, N are analogous to M_m, N_{m+1} in Section 4 above). The derivation of (29) shows that

$$\begin{aligned} L_{ma}' &\leqq 2 \max_q[(\mu(q,t) + a\,|\,\delta(q,t)\,|)]L_{ma} + d\,|\,G(t)\,|\,|\,y(t)\,|^2, \\ L_{ma}' &\geqq 2 \min_q[(\mu(q,t) - a\,|\,\delta(q,t)\,|)]L_{ma} - d\,|\,G(t)\,|\,|\,y(t)\,|^2. \end{aligned} \tag{63}$$

Similarly,

$$M' \leqq 2 \max_p[(\mu(p,t) + a\,|\,\delta(p,t)\,|)]M + d\,|\,G(t)\,|\,|\,y(t)\,|^2,$$

$$N' \geqq 2 \min_r[(\mu(r,t) - a\,|\,\delta(r,t)\,|)]N - d\,|\,G(t)\,|\,|\,y(t)\,|^2.$$

In view of (59)-(60), the number $2c$ (> 0) can play the part of both c_m and c_{m+1} in the proof of (*). The proof can now be completed along the lines of the proof of (*).

Part II. Dunkel's theorem.

11. The main theorem. Part II is concerned with a theorem of Dunkel [1] and refinements of it. In the notation of Section 2, let

$$h_* = \max_q h(q); \tag{64}$$

so that h_* is the maximum of the multiplicities of the elementary divisors of J belonging to eigenvalues λ_k satisfying $\operatorname{Re}\lambda_k = \mu^m$.

The main theorem to be proved is as follows:

(**) *Let* μ *denote the number* μ^m; h_0 *a number satisfying*

$$h_* \leqq h_0; \tag{65}$$

q_0 *a particular value of* q; k_0 *an integer satisfying* $1 \leqq k_0 \leqq h(q_0)$; *and* b *the number*

$$b = h_0 - h(q_0) + k_0, \qquad \text{(so that } k_0 \leqq b \leqq h_0). \tag{66}$$

Let $G(t)$ *be continuous on* $0 \leqq t < \infty$ *and suppose that every element* $t^{h_0-1}g_{jk\,\alpha\beta}$ *of* $t^{h_0-1}G(t)$ *is of class* $L(1,\infty)$ *or, more generally, that*

$$\int^{\infty} |\, g_{jk\,\alpha\beta} \,|\, t^{\epsilon(\alpha\beta)-\epsilon(jk)+(\epsilon(j)+\epsilon(\alpha)-1)(h_0-h(q_0))}\,dt < \infty, \tag{67}$$

where $\epsilon(jk)$ *is* k *or* 1 *according as* $j = q$ *or* $j \neq q$ *and* $\epsilon(j)$ *is* 1 *or* 0 *according as* $j = q_0$ *or* $j \neq q_0$. *Then the system* (2) *has a solution* $y = y(t)$ *with the following asymptotic properties, as* $t \to \infty$: *When* $q = q_0$,

$$y^{qk}(t) = o(e^{\mu t}t^{k-b}) \text{ if } 1 \leqq k < k_0, \tag{68_1}$$

$$y^{qk}(t) = e^{\lambda(q)t}t^{k-k_0}/(k-k_0)! + o(e^{\mu t}t^{k-b}) \text{ if } k_0 \leqq k \leqq h(q_0); \tag{68_2}$$

when $q \neq q_0$,

$$y^{qk}(t) = o(e^{\mu t}t^{k-b}) \text{ if } 1 \leqq k \leqq h(q); \tag{68_3}$$

finally, when $j = p$ *or* $j = r$,

$$y^{jk}(t) = o(e^{\mu t}t^{1-b}) \text{ if } 1 \leqq k \leqq h(j). \tag{68_4}$$

The number h_0 is not required to be an integer. Since $\epsilon(jk) = 1$ if $j \neq q$ and $1 \leqq \epsilon(jk) \leqq h(j) \leqq h_*$ if $j = q$, it is readily verified that (67) holds if

$$\int^{\infty} |\, g_{jk\,\alpha\beta} \,|\, t^{h_0-1}\,dt < \infty, \tag{67 bis}$$

or if the following pair of conditions on the "rows" of $G(t)$ is satisfied:

$$\int^{\infty} | g_{jk\,\alpha\beta}(t) | \, t^{h_0-k} dt < \infty \text{ if } j = q, \qquad \int^{\infty} | g_{jk\,\alpha\beta}(t) | \, t^{h_0-1} dt < \infty \text{ if } j \neq q$$

or if the following pair of conditions on the "columns" of $G(t)$ holds:

$$\int^{\infty} | g_{jk\,\alpha\beta} | \, t^{h_0-h(\alpha)+\beta-1} dt < \infty \text{ if } \alpha = q, \qquad \int^{\infty} | g_{jk\,\alpha\beta} | \, t^{h_0-1} dt < \infty \text{ if } \alpha \neq q.$$

After a change of variables, the case $h_0 = h_*$ of (**) is essentially Dunkel's theorem, with certain refinements. In Dunkel's theorem, assumption (67) is strengthened to (67 bis); the estimates (68_1), (68_4) and the cases $k < b$ $(= h_* - h(q_0) + k_0)$ of (68_3) are replaced by the weaker appraisal $y^{jk}(t) = o(e^{\mu t})$, as $t \rightarrow \infty$, for the values of j, k indicated; finally, (68_2) is replaced by $y^{qk}(t) \sim e^{\lambda(q)t} t^{k-k_0}/(k-k_0)!$, as $t \rightarrow \infty$, when $q = q_0$ and $k \geqq k_0$.

It will follow from the proof of (**) that, if t_0 is sufficiently large, a solution $y = y(t)$ satisfying (18) can be chosen so as to satisfy the partial set of initial conditions

$$y^{jk}(t_0) = 0 \tag{69}$$

for

$$j = p;\ j = q_0 \text{ and } k < k_0;\ j = q \neq q_0 \text{ and } k < b. \tag{70}$$

In addition to (69)-(70), one can require that $y^{q_0k}(t_0) = y_0{}^{q_0k_0} t^{k-k_0}/(k-k_0)!$ for $k_0 \leqq k < b$, in which case ($y_0{}^{q_0k_0}$ and) $y(t)$ is unique.

For the proof of (**) and for other applications, some transformations of (2) under certain changes of variables will be calculated. Before carrying out the computations involved, the change of variables to be employed will be motivated by the following remarks:

Suppose that the matrix J, which, in general, has g $(\geqq 1)$ "blocks" $J(1), \cdots, J(g)$ on its principal diagonal, consists of only one block $J(1)$. Without loss of generality, it can be supposed that the corresponding eigenvalue $\lambda(1)$ is 0. An obvious approach to theorems of type (**) is to attempt to use the variation of constants $y = e^{tJ}z$, which transforms (2) into $z' = e^{-tJ}G(t)e^{tJ}z$. But an element of the matrix $e^{-tJ}G(t)e^{tJ}$ is a linear combination of elements of $G(t)$ with coefficients which are polynomials in t, polynomials of degree n satisfying $0 \leqq n \leqq 2(h(1)-1)$. The fact that a factor as large as $t^{2(h(1)-1)}$ occurs is inconvenient. (Of course, when J consists of more than one block, even factors of the type e^{ct}, $c > 0$, might occur.)

If $J = J(1)$, then e^{tJ} has the factorization $e^{tJ} = De^{J}D^{-1}$, where D is the diagonal matrix with diagonal elements $1, t, \cdots, t^{h(1)-1}$ and e^J is a constant matrix. It turns out that a more convenient variation of constants is $y = De^Jz$. This variation of constants, when applied to the case $J = J(1)$ of (1), transforms (1) into

$$z^{11\prime} = 0, \quad tz^{12\prime} = z^{12}, \cdots, tz^{1h\prime} = (1-h)z^{1h};$$

cf. Section 13 below. If the independent variable t is replaced by a new variable defined by $ds = dt/t$, the last system becomes a system belonging to a constant diagonal matrix with distinct diagonal elements $0, -1, \cdots, 1-h$. Furthermore, $(De^J)^{-1}G(De^J)$ only involves factors as large as $t^{h(1)-1}$ (instead of the $t^{2(h(1)-1)}$ occurring in $e^{-tJ}Ge^{tJ}$).

A variation of constants of the type $y = De^Jz$ was used in [5], p. 724, for a similar purpose.

12. The transformation $y = D_1v$. The changes of variables to be used below will now be considered in detail. After the change of dependent variables $y \rightarrow e^{-\mu t}$, where $\mu = \mu^m$, it can be supposed that $\mu^m = 0$. Thus

$$\mu(j) \gtreqless 0 \text{ according as } j = p, q \text{ or } r; \tag{71}$$

cf. (4). It can then be supposed that

$$\lambda(q) = 0, \tag{72}$$

for otherwise the variable y^{qk} could be replaced by $y^{qk}e^{-\lambda(q)t}$. Under this change of (some of the) dependent variables, $g_{jk\,\alpha\beta}$ acquires factors of absolute value 1 (in fact, the factor $e^{\lambda(j)t}$ or 1 according as j is or is not a q, and the factor $e^{-\lambda(\alpha)t}$ or 1 according as α is or is not a q). Thus any assumption on the "smallness" of $|g_{jk\,\alpha\beta}|$ is not affected by this change of variables.

After these preliminaries, both the independent and dependent variables will be changed. The change of independent variables is given by

$$s = \log t, \text{ that is, } t = e^s \qquad (ds = dt/t). \tag{73}$$

It is assumed that the half-line $0 \leqq t < \infty$ has been replaced by $1 \leqq t < \infty$, which then becomes $0 \leqq s < \infty$. The change of dependent variables will be of the form $y = D_1CD_2z$, where D_1 and D_2 are diagonal matrices in which each diagonal element is a power of t to be prescribed and C is a constant matrix. In order to facilitate the computations, it will be convenient to make the changes of variables in successive steps: $y = D_1v$, $v = Cw$ and $w = D_2z$.

Let b denote a fixed number and, for each integer q, let there be given a number $b(q)$. In terms of the numbers b and $b(q)$, define $y = D_1v$ by

$$y^{qk} = t^{k-b(q)}v^{qk} \text{ and } y^{jk} = t^{1-b}v^{jk} \text{ if } j \neq q. \tag{74}$$

If a dot denotes differentiation with respect to s, then (72), (73) and (74) show that the differential system (2) becomes, when $j = q$,

$$\begin{aligned} \dot{v}^{q1} &= (b(q) - 1)v^{q1} + g_{q1\,\alpha\beta}y^{\alpha\beta}/t^{-b(q)}, \\ \dot{v}^{qk} &= (b(q) - k)v^{qk} + v^{q\,k-1} + g_{qk\,\alpha\beta}y^{\alpha\beta}/t^{k-b(q)-1}, \qquad k > 1, \end{aligned} \tag{75}$$

and, when $j \neq q$,

$$\begin{aligned} \dot{v}^{j1} &= (\lambda t - 1 + b)v^{j1} + g_{j1\,\alpha\beta}y^{\alpha\beta}/t^{-b}, \\ \dot{v}^{jk} &= (\lambda t - 1 + b)v^{jk} + tv^{j\,k-1} + g_{jk\,\alpha\beta}y^{\alpha\beta}/t^{-b}, \qquad k > 1. \end{aligned} \tag{76}$$

In (76), $\lambda = \lambda(j)$ and $\operatorname{Re}\lambda \neq 0$.

In the non-principal terms of (75)-(76), the variables $y^{\alpha\beta}$ are considered as expressed in terms of t and $v^{\alpha\beta}$ by virtue of (74) and, in addition, $t = e^s$. Let the non-principal terms of (75)-(76) be denoted by $e_{jk\,\alpha\beta}v^{\alpha\beta}$. Then, by (74) and (75)-(76), when $j = q$,

$$\begin{aligned} e_{qk\,\alpha\beta} &= g_{qk\,\alpha\beta}t^{\beta-k-b(\alpha)+b(q)+1} \text{ if } \alpha \text{ is a } q, \\ e_{qk\,\alpha\beta} &= g_{qk\,\alpha\beta}t^{1-k-b+b(q)+1} \text{ if } \alpha \text{ is not a } q, \end{aligned} \tag{77}$$

and, when $j \neq q$,

$$e_{jk\,\alpha\beta} = g_{jk\,\alpha\beta}t^{\beta-b(\alpha)+b} \text{ if } \alpha \text{ is a } q, \qquad e_{jk\,\alpha\beta} = g_{jk\,\alpha\beta}t \text{ if } \alpha \text{ is not a } q. \tag{78}$$

The matrix $\| e_{jk\,\alpha\beta} \|$ will be denoted by $E = E(s)$.

13. The transformation $v = Cw$. The constant matrix C in the variation of constants $v = Cw$ consists of g blocks $C(1), \cdots, C(g)$ on its principal diagonal and has 0 for all its other elements. $C(j)$ is an $h(j)$ by $h(j)$ matrix which is $e^{J(j)}$ or the unit $h(j)$-matrix according as $j = q$ or $j \neq q$, where $J(1), \cdots, J(g)$ are the blocks of J. Thus $v = Cw$ can be represented as

$$v^{jk} = \sum_{n=1}^{k} w^{jn}/(k-n)! \text{ or } v^{jk} = w^{jk} \tag{79}$$

according as $j = q$ or $j \neq q$. The inverse C^{-1} consists of the g blocks $C^{-1}(j)$ on its principal diagonal, where $C^{-1}(j)$ is $e^{-J(j)}$ or the unit $h(j)$-matrix according as $j = q$ or $j \neq q$.

It is readily verified that, for any scalar c,

$$\begin{bmatrix} c-1 & & & & \\ 1 & c-2 & & & \\ & 1 & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & \cdot \\ & & & 1 & c-h \end{bmatrix} C(j)$$

$$= C(j) \begin{bmatrix} c-1 & & & & \\ & c-2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & c-h \end{bmatrix},$$

where $h = h(j)$ and the matrix elements not indicated are 0. Thus the transformation (79) has the effect of diagonalizing the principal terms of (75). The system (75)-(76) therefore becomes

$$\dot{w}^{qk} = (b(q) - h)w^{qk} + h_{qk\,\alpha\beta}w^{\alpha\beta} \tag{80}$$

and, if $j \neq q$,

$$\begin{aligned} \dot{w}^{j1} &= (\lambda t - 1 - b)w^{j1} + h_{j1\,\alpha\beta}w^{\alpha\beta}, \\ \dot{w}^{jk} &= (\lambda t - 1 - b)w^{jk} + tw^{j\,k-1} + h_{jk\,\alpha\beta}w^{\alpha\beta}, \qquad k > 1, \end{aligned} \tag{81}$$

where $H = \| h_{jk\,\alpha\beta} \|$ is given by $H = C^{-1}EC$.

It is readily verified from (79) and the description of C^{-1} that

$$\begin{aligned} h_{jk\,\alpha\beta} &= e_{jk\,\alpha\beta} \text{ if neither } j \text{ nor } \alpha \text{ is a } q, \\ h_{jk\,\alpha\beta} &= \sum_{\gamma=\beta}^{h(\alpha)} e_{jk\,\alpha\gamma}/(\gamma-\beta)\,! \text{ if } j \text{ is not, but } \alpha \text{ is, a } q, \\ h_{jk\,\alpha\beta} &= \sum_{n=1}^{k} (-1)^{k-n} e_{jk\,\alpha\beta}/(k-n)\,! \text{ if } j \text{ is, but } \alpha \text{ is not, a } q, \\ h_{jk\,\alpha\beta} &= \sum_{n=1}^{k} \sum_{\gamma=\beta}^{h(\alpha)} (-1)^{k-n} e_{jn\,\alpha\beta}/(k-n)\,!(\gamma-\beta)\,! \text{ if both } j \text{ and } \alpha \text{ are } q\text{'s}. \end{aligned} \tag{82}$$

It can be noted that $h_{jk\,\alpha\beta}$ is a linear combination, with constant coefficients, of the elements $e_{jn\,\alpha\gamma}$, where j and α are fixed and n $(\leqq k)$, γ $(\geqq \beta)$ vary.

14. The transformation $w = D_2z$. The last change of variables $w = D_2z$ is given by

$$w^{jk} = t^c z^{jk} \text{ if } j = q_0 \text{ and } k \neq k_0, \qquad w^{jk} = z^{jk} \text{ otherwise,} \tag{83}$$

where c is a constant and $t = e^s$. This is the identity transformation if $c = 0$. The system (80)-(81) is transformed by (83) into

$$\dot z^{q_0k} = (b(q_0) - c - k)z^{q_0k} + b_{q_0k\,\alpha\beta}z^{\alpha\beta} \text{ if } k \neq k_0,$$

(84)

$$\dot z^{qk} = (b(q) - k)z^{qk} + b_{qk\,\alpha\beta}z^{\alpha\beta} \text{ if } q \neq q_0 \text{ or if } (qk) = (q_0k_0),$$

and, if $j \neq q$,

$$\dot z^{j1} = (\lambda t - 1 - b)z^{j1} + b_{j1\,\alpha\beta}z^{\alpha\beta},$$

(85)

$$\dot z^{jk} = (\lambda t - 1 - b)z^{jk} + tz^{j\,k-1} + b_{jk\,\alpha\beta}z^{\alpha\beta} \text{ if } k > 1,$$

where $B = \| b_{jk\,\alpha\beta} \|$ is the matrix $D_2^{-1}HD_2$.

Since (83) is $w = D_2 z$, it is easily verified that, if $k \neq k_0$,

$$b_{q_0k\,\alpha\beta} = h_{q_0k\,\alpha\beta} \text{ if } \alpha = q_0, \beta \neq k_0,$$

(86_1)

$$b_{q_0k\,\alpha\beta} = h_{q_0k\,\alpha\beta}t^{-c} \text{ if } \alpha \neq q_0 \text{ or if } (\alpha\beta) = (q_0k_0);$$

and, if $j \neq q_0$ or if $(jk) = (q_0k_0)$,

$$b_{jk\,\alpha\beta} = h_{jk\,\alpha\beta}t^{c} \text{ if } \alpha = q,\ \beta \neq k_0,$$

(86_2)

$$b_{jk\,\alpha\beta} = h_{jk\,\alpha\beta} \text{ if } \alpha = q \text{ or if } (\alpha\beta) = (q_0k_0).$$

15. Estimates for *B*. The proof of (**) now proceeds as follows: Assume the normalizations (71) and (72). Introduce the change of variables (83) and the successive variations of constants (74), (79) and (83). In (84), the numbers $b(q)$ and b will be chosen as follows:

(87) $$b(q_0) = k_0; \quad b(q) = b \text{ if } q \neq q_0; \quad b = h_0 - h(q_0) - k_0$$

(this number b is the same as in (66)). In (83), let

(88) $$c = h(q_0) - h_0 \qquad \text{(so that } c \leqq 0).$$

It will be verified that

(89) $$\int^{\infty} | b_{jk\,\alpha\beta} |\, ds < \infty \text{ for arbitrary } (jk\,\alpha\beta)$$

and, in addition, that

(90) $$\int^{\infty} e^{-cs} | b_{jk\,\alpha\beta} |\, ds < \infty \text{ if } (jk) = (q_0k_0), \qquad (e^{-cs} = t^{h_0 - h(q_0)}).$$

To this end, the matrices E, H and B will be examined in turn. In view of (87) and the relations (77)-(78), it is seen that $e_{q_0k\,\alpha\beta}$ is $g_{q_0k\,\alpha\beta}$ times

$t^{\beta-k+1}$, $t^{\beta-k-b+k_0+1}$ or $t^{1-k-b+k_0+1}$ according as $\alpha = q_0$, $\alpha = q \neq q_0$ or $\alpha \neq q$. In view of the definitions of b in (66) and of $\epsilon(jk)$, $\epsilon(j)$ following (67), it is seen that

$$| e_{jk\,\alpha\beta} | = | g_{jk\,\alpha\beta} | \, t^{\epsilon(\alpha\beta)-\epsilon(jk)+(\epsilon(\alpha)-1)(h_0-h(q_0))+1} \tag{91}$$

if $j = q_0$. If $q \neq q_0$, then $e_{qk\,\alpha\beta}$ is $g_{qk\,\alpha\beta}$ times $t^{\beta-k-k_0+b+1}$, $t^{\beta-k+1}$ or t^{1-k+1} according as $\alpha = q_0$, $\alpha = q \neq q_0$ or $\alpha \neq q$. Thus (91) holds if $j = q \neq q_0$. Finally, if $j \neq q$, then $e_{jk\,\alpha\beta}$ is $g_{jk\,\alpha\beta}$ times $t^{\beta-k_0+b}$, t^{β} or t according as $\alpha = q_0$, $\alpha = q \neq q_0$ or $\alpha \neq q$. In this case, $j \neq q$, the relation (91) also holds.

Since $ds = dt/t$, it follows from (67) and $\epsilon(j)(h_0 - h(q_0)) \geqq 0$ that

$$\int^{\infty} | e_{jk\,\alpha\beta} | \, ds < \infty, \tag{92}$$

and, from $\epsilon(j)(h_0 - h(q_0)) = h_0 - h(q_0)$ where $j = q_0$, that

$$\int^{\infty} e^{(h_0-h(q_0))s} \, | e_{q_0k\,\alpha\beta} | \, ds < \infty \qquad (t = e^s). \tag{93}$$

By (82), the relations corresponding to (92), (93) hold if $e_{jk\,\alpha\beta}$ is replaced by $h_{jk\,\alpha\beta}$. Finally, (89) and (90) follow from these relations, (86) and (88), since $t = e^s$.

16. Completion of the proof of (**). Note that $b(q_0) - c = b$, by (87) and (88), so that the set of differential equations (84) can be rewritten as

$$\dot{z}^{qk} = b_{qk\,\alpha\beta} z^{\alpha\beta} \tag{94}$$

if $(qk) = (q_0k_0)$, and as

$$\dot{z}^{qk} = (b - k) z^{qk} + b_{qk\,\alpha\beta} z^{\alpha\beta} \tag{95}$$

if $(qk) \neq (q_0k_0)$.

Let (q^*k^*) denote either (q_0k_0) or any pair of indices (qk) such that $q \neq q_0$ and $(h(q) \leqq) k = b$. (For example, if h_0 (and, therefore, b) is not an integer, then the only pair (q^*k^*) is (q_0k_0).) From the definition of (q^*k^*), the differential equation for z^{jk} has the form (94) if and only if $(jk) = (q^*k^*)$. Put

$$L = L(s) = \Sigma\Sigma \, | z^{q^*k^*} |^2. \tag{96}$$

Note that $|(\mathrm{Re}\,\lambda(j) \pm a)t - 1 + b| \geqq \text{const.} > 0$ for large t if $a\ (> 0)$ is sufficiently small. Thus (III) in Section 10 implies that, for a sufficiently large S and a given set of numbers $z_0^{q^*k^*}$, not all 0, there exists a unique solution $z = z(s)$ of (84)-(85) satisfying the partial set of initial conditions

$$z^{jk}(S) = 0 \text{ if } j = p;\ j = q \text{ and } k < b \text{ but } (jk) \neq (q_0k_0), \tag{97}$$
$$z^{q^*k^*}(S) = z_0^{q^*k^*}$$

and, as $s \to \infty$, the relation

$$z^{jk} = o(L^{\frac{1}{2}}(s)) \text{ if } (jk) \neq (q^*k^*). \tag{98}$$

It follows from (94), (96) and (98) that

$$|\dot{L}| \leqq \text{const.}\, |B(s)|\, L. \tag{99}$$

Consequently (89) implies that, as $s \to \infty$,

$$L(s) = O(1) \text{ and } 1/L(s) = O(1). \tag{100}$$

In particular, $z^{jk}(s) = O(1)$ as $s \to \infty$.

Thus, by (89) and (94),

$$\lim_{s \to \infty} z^{q^*k^*}(s) = z^{q^*k^*}(\infty) \text{ exists.} \tag{101}$$

In particular, $L(\infty)$ exists and $L(\infty) \neq 0$, by the second part of (100). It follows from the superposition principle that if S is fixed and sufficiently large, then to any preassigned set of numbers $z^{q^*k^*}(\infty)$, not all 0, there belongs a unique set of numbers $z_0^{q^*k^*}$, not all zero, with the property that the solution satisfying (97) and (98) also satisfies (101).

Let $z^{q^*k^*}(\infty)$ be 1 or 0 according as $(q^*k^*) = (q_0k_0)$ or $(q^*k^*) \neq (q_0k_0)$. Thus, by (97) and (101), the corresponding solution satisfies, as $s \to \infty$,

$$z^{jk} \to 1 \text{ or } z^{jk} = o(1) \tag{102}$$

according as $(jk) = (q_0k_0)$ or $(jk) \neq (q_0k_0)$. In view of (90), (94) and (101), the relation

$$z^{jk}(s) = 1 + \int_s^\infty b_{jk\,\alpha\beta}(u) z^{\alpha\beta}(u)\,du, \tag{103}$$

where $(jk) = (q_0k_0)$, gives the estimate

$$z^{jk} = 1 + o(t^{h(q_0)-h_0}) \text{ if } (jk) = (q_0k_0) \text{ and } t = e^s. \tag{104}$$

By (83) and (88), the corresponding solution $w = w(s)$ of (80)-(81) satisfies, as $s \to \infty$,

$$w^{jk} = o(t^{h(q_0)-h_0}) \text{ if } j = q_0,\ k \neq k_0,$$
$$w^{jk} = 1 + o(t^{h(q_0)-h_0}) \text{ if } j = q_0,\ k = k_0,$$
$$w^{jk} = o(1) \text{ if } j \neq q_0.$$

Thus, by (79), the corresponding vector $v = v(s)$ is

$$v^{jk} = o(t^{h(q_0)-h_0}) \text{ if } j = q_0,\ k < k_0,$$

$$v^{jk} = 1/(k - k_0)! + o(t^{h(q_0)-h_0}) \text{ if } j = q_0,\ k \geqq k_0,$$

$$v^{jk} = o(1) \text{ if } j \neq q_0.$$

Finally, (74) and (87) imply that the corresponding solution $y = y(t)$ of (2) satisfies (68), when the normalization (72) is assumed. This proves (**).

17. An application of (**). This application concerns the asymptotic integration of a linear system of differential equations, in which the coefficient matrix is the sum of a periodic matrix $P(t)$ and of a matrix $F(t)$ which is "small" for large t.

(†) *Let* $P(t)$ *be a continuous* d *by* d *periodic matrix of period* 1,

$$P(t+1) = P(t). \tag{105}$$

Let $F(t)$ *be continuous on* $0 \leqq t < \infty$ *and satisfy*

$$\int_0^\infty t^{2(d-1)} |F(t)|\, dt < \infty. \tag{106}$$

Let $x = x_0(t) \not\equiv 0$ *be a solution of*

$$x' = P(t)x \tag{107}$$

and let $\mu = \lim t^{-1} \log |x_0(t)|$, *as* $t \to \infty$. *Then*

$$x' = (P(t) + F(t))x \tag{108}$$

has a solution $x = x(t)$ *satisfying*

$$x(t) - x_0(t) = o(e^{\mu t}) \text{ as } t \to \infty. \tag{109}$$

Standard theorems state that the system (107) has a fundamental matrix $X = X(t)$ of the form

$$X(t) = Z(t)e^{tJ}, \text{ where } Z(t+1) = Z(t) \tag{110}$$

and J is a constant matrix. This makes it clear that the limit μ exists for any solution $x = x_0(t) \not\equiv 0$ of (107) and $\mu = \operatorname{Re} \lambda$ holds for some characteristic number λ of J.

It will be clear from the proof of (†) that condition (106) can be relaxed and assertion (109) strengthened if the solutions of (107), say the matrix (110), are known. Condition (106), for example, can be weakened by replacing the $d-1$ in (106) by h_*-1, where h_* is the maximum of the multiplicities of the elementary divisors of J belonging to characteristic numbers λ satisfying $\operatorname{Re}\lambda=\mu$.

The proof of (†) will depend on the fact that the variation of constants

$$x=Z(t)y \tag{111}$$

transforms (107) into $y'=Jy$ and (108) into $y'=(J+G(t))y$, where

$$G(t)=Z^{-1}(t)F(t)Z(t). \tag{112}$$

It is sufficient to verify this transformation of (108), since (107) is the case $F(t)\equiv 0$ of (108). If (111) is substituted into (108), then the resulting differential equation for y is

$$y'=(Z^{-1}(P+F)Z-Z^{-1}Z')y.$$

Differentiation of (110) gives $Z'e^{tJ}+ZJe^{tJ}=PZe^{tJ}$, since $(e^{tJ})'=Je^{tJ}$ and $X(t)$ is a fundamental matrix of (108). Hence $Z'=PZ-ZJ$. If this is substituted into the last formula line, there results the equation $y'=(J+G(t))y$, where $G(t)$ is given by (112).

Since $Z(t)$ is periodic and non-singular, the matrices $Z(t)$ and $Z^{-1}(t)$ are $O(1)$ as $t\to\infty$. Hence (106) holds if $F(t)$ is replaced by $G(t)$. Thus, according to (**), to any solution $y=y_0(t)\not\equiv 0$ of $y'=Jy$, there corresponds at least one solution $y=y(t)$ of $y'=(J+G)y$ satisfying $y-y_0=o(e^{\mu t})$, as $t\to\infty$, where $\mu=\lim t^{-1}\log|y_0(t)|$.

The assertion of (†) now follows by choosing $y_0(t)$ to be $Z^{-1}(t)x_0(t)$ and $x(t)$ to be $Z(t)y(t)$.

Part III. Variants and applications.

18. Statement of results. The first portion of Part III concerns asymptotic integrations of (2) under conditions different from those occurring in (**). The first theorem to be proved is as follows:

(i) *Let* $\lambda_1,\cdots,\lambda_d$ *be the eigenvalues of the coefficient matrix* J *in* (1) *and let one of these, say* $\lambda=\lambda_m$, *satisfy*

$$\mu_m\neq\mu_j \text{ for } j\neq m, \qquad (\mu_k=\operatorname{Re}\lambda_k). \tag{113}$$

Let $G(t)$ *be of class* L^γ *for some* γ *on the range* $1\leqq\gamma\leqq 2$, *that is, let*

$$\int^{\infty}|G(t)|^\gamma\,dt<\infty \text{ for some } \gamma, \text{ where } 1\leqq\gamma\leqq 2. \tag{114}$$

Then (2) *possesses a solution* $y = y(t) = (y^1(t), \cdots, y^d(t))$ *the components of which satisfy, as* $t \to \infty$,

$$(115_1) \qquad y^m(t) \sim \exp \int^t (\lambda_m + g_{mm})\,du,$$

$$(115_2) \qquad y^j(t) = o(|\,y^m(t)\,|) \text{ if } j \neq m.$$

In (115_1), $g_{mm} = g_{mm}(t)$ is the m-th diagonal element of $G(t)$. Assumption (113) implies that the (single) elementary divisor of J for which $\operatorname{Re} \lambda(j) = \mu_m$ is simple. Accordingly, the proof of (i) does not depend on the changes of variables discussed in Sections 12-14 above. The proof of (i) will be given in Section 21 below.

(i) is a generalization of Theorem (VII) in [2], p. 575, dealing with a second order equation $x'' + fx = 0$, where $f = f(t)$ is continuous, real-valued, and $f(t) + 1$ is of class L^γ on $0 \leqq t < \infty$.

Assertion (i) remains valid if the eigenvalues λ_j of J are not constants but are continuous functions $\lambda_j = \lambda_j(t)$ of t, subject to

$$(116) \qquad |\,\mu_j(t) - \mu_m(t)\,| \geqq \text{const.} > 0 \text{ if } j \neq m \qquad (\mu_j = \operatorname{Re} \lambda_j)$$

for large t.

In particular, if $J = J(t)$ is a diagonal matrix with diagonal elements $\lambda_j = \lambda_j(t)$ satisfying (116) for every pair of indices j, m $(\neq j)$, then (i) furnishes the asymptotic behavior of a fundamental set of solutions of (2).

There exists an analogue of (i) which involves no assumption on the elementary divisors of J, but imposes more severe conditions on $G(t)$. This analogue does not contain (i).

(ii) *Let* h_* *be defined by* (64); q_0 *a particular value of* q; k_0 *an integer satisfying* $1 \leqq k \leqq h(q_0)$. *Let* h_0 *be a number satisfing* (65) *and let*

$$(117) \qquad b - k \neq 0 \text{ if } 1 \leqq k \leqq h(q) \text{ and } q \neq q_0,$$

where b *is given by* (66). *For some* γ, *where* $1 \leqq \gamma \leqq 2$, *let*

$$(118) \qquad \int^\infty t^{\gamma h_0 - 1}\,|\,G(t)\,|^\gamma\,dt < \infty$$

or, more generally, for every set of indices $(jk\,\alpha\beta)$, *let*

$$(119) \qquad \int^\infty |\,g_{jk\,\alpha\beta}\,|^\gamma\, t^{\gamma\{\epsilon(\alpha\beta) - \epsilon(jk) + (\epsilon(j) + \epsilon(\alpha) - 1)(h_0 - h(q_0))\} + \gamma - 1}\,dt < \infty,$$

where $\epsilon(jk)$ *and* $\epsilon(j)$ *are defined as after* (67). *Then* (2) *has a solution* $y = y(t)$ *with the following asymptotic properties, as* $t \to \infty$: *When* $q = q_0$,

$$(120_1) \qquad y^{qk} = o(e^{\mu t} t^{k-b} \mid \exp \mid) \text{ if } k < k_0,$$

$$(120_2) \qquad y^{qk} = e^{\lambda(q)t} t^{k-k_0} \exp/(k-k_0)\,! + o(e^{\mu t} t^{k-b} \mid \exp \mid) \text{ if } k \geqq k_0;$$

when $q \neq q_0$,

$$(120_3) \qquad y^{qk} = o(e^{\mu t} t^{k-b} \mid \exp \mid);$$

finally, when $j = p$ *or* $j = r$,

$$(120_4) \qquad y^{jk} = o(e^{\mu t} t^{1-b} \mid \exp \mid),$$

where

$$(121) \quad \exp = \exp\{ \int^{t} \sum_{k=1}^{k_0} \sum_{\beta=k_0}^{h(q_0)} (-1)^{k_0-k} u^{\beta-k} g_{q_0 k\, q_0 \beta}(u)\, du/(k_0 - k)\,!\,(\beta - k_0)\,! \}.$$

The condition (117) is vacuous if, for example, there is no q distinct from q_0, that is, if $\operatorname{Re}\lambda(j) \neq \operatorname{Re}\lambda(q)$ when $j \neq q$. The condition (117) holds when h_0 (hence, b) is not an integer.

19. The differential inequalities (37), (44). The proof of (i) will depend on a sharpened form of the differential inequalities (37), (44), occurring in the proof of (*), and on a theorem on Riccati differential inequalities. Assertion (ii) will be reduced essentially to (i) by the changes of variables used for similar purposes in Part II.

In order to obtain the desired differential inequalities, note that the expression $\mid G(t) \mid \cdot \mid y(t) \mid^2$ occurring (twice) in the line following (28) can be replaced by $\mid G \mid \cdot \mid y \mid \cdot \mid y^{jk} \mid$. Thus the derivation of (29) gives

$$\mid L_{ma}' - 2\mu^m L_{ma} \mid \leqq 2aL_{ma} + 2 \mid y \mid \cdot \mid G \mid \cdot \sum_q \sum_k a^{2k} \mid y^{qk} \mid.$$

Since $a^{2k} \mid y^{qk} \mid \leqq a^k \mid y^{qk} \mid \leqq L_{ma}^{\frac{1}{2}}$, the last double sum does not exceed $dL_{ma}^{\frac{1}{2}}$, and so

$$(122_m) \qquad \mid L_{ma}' - 2\mu^m L_{ma} \mid \leqq 2aL_{ma} + 2d \mid y \mid \cdot \mid G \mid L_{ma}^{\frac{1}{2}}$$

can replace (29). This leads to the analogues

$$(123_m) \qquad M_m' \leqq 2(\mu^{m-1} + a) M_m + 2d^2 \mid y \mid \cdot \mid G \mid M_m^{\frac{1}{2}},$$

$$(124_{m+1}) \qquad N_{m+1}' \geqq 2(\mu^{m+1} - a) N_{m+1} - 2d^2 \mid y \mid \cdot \mid G \mid N_{m+1}^{\frac{1}{2}}$$

of (33_m) and (34_{m+1}).

If $L_{ma}(t) \neq 0$ for some t-value (hence for all nearby t-values), then (122_m) can be divided by $2L_{ma}^{\frac{1}{2}}$ to give

$$|(L_{ma}^{\frac{1}{2}})' - \mu^m L_{ma}^{\frac{1}{2}}| \leqq aL_{ma}^{\frac{1}{2}} + d\,|y|\cdot|G|. \tag{125_m}$$

Corresponding to the assumptions $M_m(t) \neq 0$ or $N_{m+1}(t) \neq 0$, (123_m), (124_{m+1}) become, respectively,

$$(M_m^{\frac{1}{2}})' \leqq (\mu^{m-1} + a)M_m^{\frac{1}{2}} + d^2\,|y|\cdot|G|, \tag{126_m}$$

$$(N_{m+1}^{\frac{1}{2}})' \geqq (\mu^{m+1} - a)N_{m+1}^{\frac{1}{2}} - d^2\,|y|\cdot|G|. \tag{127_{m+1}}$$

These differential inequalities make it possible to replace the auxiliary functions (35) and (42) of Sections 4 and 6, respectively, by

$$v(t) = v_m(t) = N_m^{\frac{1}{2}}/(N_m^{\frac{1}{2}} + M_m^{\frac{1}{2}}), \tag{128}$$

$$w(t) = v_{m+1}(t) = N_{m+1}^{\frac{1}{2}}/(N_{m+1}^{\frac{1}{2}} + M_{m+1}^{\frac{1}{2}}). \tag{129}$$

If the solution $y = y(t)$ of (2) determining L_{ma}, M_m, N_m is not identically 0, then (128) and (129) are defined, continuous, and satisfy

(130_1) $\quad 0 \leqq v \leqq 1;$ $\qquad$ (130_2) $\quad 0 \leqq w \leqq 1.$

If $|y|$ is replaced by its majorant $a^{-d}(M_m^{\frac{1}{2}} + N_m^{\frac{1}{2}})$ in (125)-(127), then the derivation of (37) in Section 4 shows that (128) satisfies

$$v' \geqq \tfrac{1}{2}c_m v(1 - v) - d^2 a^{-d}\,|G|, \qquad c_m > 0, \tag{131}$$

provided that $N_m \neq 0$ and $M_m \neq 0$. Similarly,

$$w' \geqq \tfrac{1}{2}c_{m+1} w(1 - w) - d^2 a^{-d}\,|G|, \qquad c_{m+1} > 0, \tag{132}$$

provided that $N_{m+1} \neq 0$ and $M_{m+1} \neq 0$.

The considerations below will deal with solutions $y = y(t)$ ($\neq 0$) satisfying, as $t \to \infty$,

(133_1) $\quad v(t) \to 1,$ $\qquad$ (133_2) $\quad w(t) \to 0;$

so that

$$M_m = o(L_{ma}) \text{ and } N_{m+1} = o(L_{ma}) \text{ as } t \to \infty. \tag{134}$$

Thus, for large t, $L_{ma} \neq 0$, and so $N_m \neq 0$. For large t, $M_m = 0$ if and only if $v = 1$, while $N_{m+1} = 0$ if and only if $w = 0$. Consequently, for large t, (131) and (132) hold except possibly at those t-values at which $v = 1$ and $w = 0$, respectively.

20. A lemma on Riccati inequalities. Assertion (i) will be obtained from (131), (132) and the following theorem on a Riccati differential inequality:

LEMMA 3. *Let $f=f(t)$ be a real-valued, continuous function on $0 \leqq t < \infty$ and let $f(t)$ be of class L^{γ} for some γ, where $1 \leqq \gamma \leqq 2$. Let $z=z(t)$ be a real-valued continuous function for large t satisfying*

$$(135) \qquad 0 \leqq z \leqq 1$$

and either

$$(136_1) \qquad z \to 1 \qquad \text{or} \qquad (136_2) \qquad z \to 0$$

as $t \to \infty$. In addition, let $z(t)$ have, on the respective open t-sets $\{z(t) \neq 1\}$, $\{z(t) \neq 0\}$, a continuous derivative satisfying a differential inequality of the form

$$(137) \qquad z' \geqq cz(1-z)+f, \qquad \text{where } c = \text{const.} > 0.$$

Then, corresponding to the alternatives in (136),

$$(138_1) \quad \int^{\infty} |z-1|^2\,dt < \infty \qquad \text{or} \qquad (138_2) \quad \int^{\infty} |z|^2\,dt < \infty.$$

This lemma is an analogue of (III) in [2], pp. 568-569. As in [2], the assertions "$z-1$ is of class L^2 or z is of class L^2" can be improved to "$z-1$ is of class L^{γ} or z is of class L^{γ}."

In order to prove Lemma 3, consider the case (136_2) (the other case can be treated similarly). If $[U, V]$ is an interval on which $z > 0$, then (137) is valid on this interval. Multiplication by $z > 0$ does not affect the inequality, and so a quadrature leads to

$$(139) \qquad c\int_U^V z^2(1-z)\,dt \leqq \tfrac{1}{2}z^2(V) - \tfrac{1}{2}z^2(U) + \int_U^V |fz|\,dt.$$

By continuity, (139) holds if $z(V)=z(U)=0$; that is,

$$c\int_U^V z^2(1-z)\,dt \leqq \int_U^V |fz|\,dt.$$

This inequality remains valid if the domain of integration (U, V) is replaced by a sum of open intervals bounded by zeros of $z(t)$ and on which $z > 0$. Since the integral of $z^2(1-z)$ on a t-set where $z=0$ is 0, it follows that the last inequality is valid whenever $z(V)=z(U)=0$ and $z(t) \geqq 0$ on (U, V). Hence (139) is valid for arbitrary intervals $[U, V]$.

If U is so large that $0 \leqq z(t) \leqq \frac{1}{2}$ for $t \geqq U$, then (139) and Hölder's inequality give

$$\tfrac{1}{2}c \int_U^V z^2 dt \leqq \tfrac{1}{2} + \Big(\int_U^V |f|^\gamma dt\Big)^{1/\gamma} \Big(\int_U^V |z|^{\gamma/(\gamma-1)} dt\Big)^{1-1/\gamma}.$$

If $\gamma = 1$, then the last integral does not occur and (138_2) follows from the fact that f is of class $L^\gamma = L$. If $\gamma > 1$, then (136_2) implies that $|z|^{\gamma/(\gamma-1)} \leqq z^2$ for large t, and so (138_2) follows from the fact that $1 - 1/\gamma < 1$. This proves Lemma 3.

21. Proof of (i). In order to prove (i), let $y = y(t)$ be a solution of (2) satisfying (115_2), as supplied by (*); cf. (II*) and the remarks following it. To this solution, there belong auxiliary functions $v(t)$ and $w(t)$, given by (128) and (129), which are defined and continuous for large t, satisfy the limit relations (133), and possess continuous derivatives satisfying the differential inequalities (131), (132) on the respective t-sets $\{v(t) \neq 1\}$, $\{w(t) \neq 0\}$.

Condition (114) of (i) and assertion (138) in Lemma 3 imply that

$$\int^\infty w^2 dt < \infty \quad \text{and} \quad \int^\infty (1 - v)^2 dt < \infty.$$

From the definitions of v and w, the functions $(M_m/L_{ma})^{\frac{1}{2}}$ and $(N_{m+1}/L_{ma})^{\frac{1}{2}}$ are of class L^2; cf. (134). Since, by condition (113) in (i), there is only one elementary divisor belonging to a λ_j for which $\operatorname{Re} \lambda_j = \mu_m$ and this one is simple, it follows that $L_{ma} = a^2 |y^m|^2$. Hence, y^j/y^m is of class L^2 if $j \neq m$. If $\gamma > 1$, so that $\gamma/(\gamma - 1) > 2$, then (115_2) implies that y^j/y^m is of class $L^{\gamma/(\gamma-1)}$ if $j \neq m$.

The differential equation for y^m in (2) is of the form (15), where $G = \| g_{mj} \|$. Since, for large t, y^m does not vanish,

$$y^{m\prime}/y^m = (\lambda_m + g_{mm}) + \sum_{j \neq m} g_{mj}(y^j/y^m).$$

Thus a quadrature gives

$$y^m(t)/\exp \int^t (\lambda_m + g_{mm}) du = \exp \int^t \sum_{j \neq m} g_{mj}(y^j/y^m) du.$$

If $\gamma = 1$, then, by (114) and (115_2), $g_{mj}(y^j/y^m)$ is of class L; the same holds if $\gamma > 1$, since y^j/y^m is of class $L^{\gamma/(\gamma-1)}$. Thus the expression on the right side of the last formula line tends, as $t \to \infty$, to a non-vanishing constant. If

$y(t)$ is replaced by const. $y(t)$, with a suitable choice of const. $\neq 0$, then (115_1) holds. This proves (i).

22. Proof of (ii). In order to prove (ii), assume the normalizations (72). Introduce the change of independent variables (73) and the successive variations of constants (74), (79) and (83), in which (87) and (88) hold. This transforms (2) into the system (84)-(85), where a differential equation in (84) can be written as (94) or (95) according as $(qk) = (q_0k_0)$ or $(qk) \neq (q_0k_0)$. The assumption (117) of (ii) means that, in the notation of Section 16 above, there is only one pair of indices (q^*k^*), namely, (q_0k_0).

The identities (91) along with the relations (82) and (87) show that the matrix $B = B(s)$ in (84)-(85) is of class L^γ (as a function of s), by virtue of (119). If (III) is used instead of $(^*)$ or (II^*), then the proof of (i) shows that (84)-(85) has a solution $z = z(s)$ satisfying, as $s \to \infty$,

$$(140_1) \qquad z^{q_0k_0} \sim \exp \int^s b_{q_0k_0\ q_0k_0} ds,$$

$$(140_2) \qquad z^{jk} = o(|z^{q_0k_0}|) \quad \text{if} \quad (jk) \neq (q_0k_0).$$

The assumption (119), the identities (91) and the relations (82), (86) show that $t^{h_0-h(q_0)} b_{q_0k_0\ \alpha\beta}$, where $(\alpha\beta) \neq (q_0k_0)$, is of class L^γ (as a function of s). Since the case $(jk) = (q_0k_0)$ of (94) can be written as

$$(\log z^{q_0k_0})' = b_{q_0k_0\ q_0k_0} + \sum_{(\alpha\beta)\neq(q_0k_0)} \sum b_{q_0k_0\ \alpha\beta}(z^{\alpha\beta}/z^{q_0k_0}),$$

the derivation of (140_1) shows that (140_1) can be improved to

$$(140_1\ \text{bis}) \qquad z^{q_0k_0} = (1 + o(t^{h(q_0)-h_0})) \exp \int^s b_{q_0k_0\ q_0k_0} ds,$$

as $s \to \infty$.

In view of (77)-(78), (82) and (86),

$$b_{q_0k_0\ q_0k_0} = \sum_{n=1}^{k_0} \sum_{\beta=k}^{h(q_0)} (-1)^{k_0-n} t^{\beta-n+1}\ g_{q_0n\ q_0\beta}/(k_0-n)!(\beta-k_0)!.$$

Thus, the formulae (120) follow from (73) and (74), (79), (83). This proves (ii).

23. Applications to a second order equation. Let x be a scalar variable. It was shown in [2], pp. 570-573, that if, in the differential equation

$$(141) \qquad x'' - (\lambda + \phi(t))x = 0,$$

λ is a positive constant and $\phi(t)$ is real-valued and continuous for large t, then necessary and sufficient for (141) to be non-oscillatory and to possess a pair of solutions satisfying

$$x'/x \to \pm\lambda^{\frac{1}{2}} \quad \text{as} \quad t \to \infty \tag{142}$$

is that

$$\underset{0 \leqq V < \infty}{\text{l. u. b.}} \left| \int_U^{U+V} \phi(t)\,dt \right| / (1+V) \to 0 \quad \text{as} \quad U \to \infty. \tag{143}$$

The main theorem (*) of Part II gives an immediate extension of one half of this theorem to the case when λ is a complex number, but is not negative or 0, and $\phi(t)$ is allowed to be complex-valued. (For an analogous theorem when λ is a negative number and $\phi(t)$ is real-valued, cf. [4].)

(α) *Let* λ *be a complex number such that*

$$\operatorname{Re} \lambda^{\frac{1}{2}} \neq 0 \tag{144}$$

and let $\phi(t)$ *be a continuous complex-valued function satisfying*

$$\underset{0 \leqq V < \infty}{\text{l. u. b.}} \int_U^{U+V} |\phi(t)|\,dt / (1+V) \to 0 \quad \text{as} \quad U \to \infty. \tag{145}$$

Then (141) *has a pair of solutions which do not vanish for large* t *and satisfy* (142).

In order to prove (α), write (141) as a first order system $y' = (J + G(t))y$ for the vector $y = (y^1, y^2)$, where

$$y^1 = x' + \lambda^{\frac{1}{2}}x, \qquad y^2 = x' - \lambda^{\frac{1}{2}}x, \tag{146}$$

for some fixed choice of the square root $\lambda^{\frac{1}{2}}$. It is readily verified that (146) satisfies $y' = (J + G(t))y$, where

$$J = \begin{pmatrix} \lambda^{\frac{1}{2}} & 0 \\ 0 & -\lambda^{\frac{1}{2}} \end{pmatrix} \text{ and } G(t) = \tfrac{1}{2}\lambda^{-\frac{1}{2}}\phi \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}. \tag{147}$$

Conditions (144), (145) imply conditions (54), (6) of (II*), respectively. Hence (II*) implies the existence of a solution $y = (y^1(t), y^2(t))$ satisfying

$$y^1(t) = o(|\,y^2(t)\,|) \quad \text{as} \quad t \to \infty. \tag{148}$$

The corresponding solution $x = x(t)$ of (141) satisfies, by (146),

$$(1 + o(1))x' = (\lambda^{\frac{1}{2}} + o(1))x \quad \text{as} \quad t \to \infty;$$

that is, $x'/x \to \lambda^{\frac{1}{2}}$. Since the choice of $\lambda^{\frac{1}{2}}$ is arbitrary, assertion (α) follows.

Theorem (VII) in [2], p. 575, which deals with the case $\lambda = 1$ and a real-valued $\phi = \phi(t)$, can be extended as follows:

(β) *If* (144) *holds and* (145) *is strengthened to*

$$(149) \qquad \int^{\infty} |\phi(t)|^{\gamma}\, dt < \infty \text{ for some } \gamma, \text{ where } 1 \leqq \gamma \leqq 2,$$

then (141) *has a pair of solutions satisfying* (142) *and*

$$(150) \qquad x \sim \exp \pm \int^{t} (\lambda^{\frac{1}{2}} + \tfrac{1}{2}\lambda^{-\frac{1}{2}} \phi(u))\, du \quad \text{as} \quad t \to \infty.$$

If, in addition, $\phi(t)$ is (improperly) integrable at $t = \infty$, that is, if

$$\int^{\infty} \phi(t)\, dt = \lim_{T \to \infty} \int^{T} \phi(t)\, dt \text{ exists,}$$

then (150) can be replaced by $x \sim \exp(\pm \lambda^{\frac{1}{2}} t)$.

The assertions of (β) follow from (i), since (i) supplies the existence of a vector (146) satisfying (148) and

$$y^1 \sim \text{const.} \exp \int^{t} (\lambda^{\frac{1}{2}} + \tfrac{1}{2}\lambda^{-\frac{1}{2}} \phi(u))\, du,$$

where const. $\neq 0$ can be chosen to be $2\lambda^{\frac{1}{2}}$.

Cf. (vi) and (vii) in [6], p. 722, for the analogues of (α) and (β) when $\lambda = 0$. It is clear from (α) and (β) that the theorems (vi), (vii) in [6] have analogues when the coefficient function $f(t)$ there is complex-valued.

24. Applications to a second order equation (continued). There will now be considered theorems of the type (α), (β) in which (141) is replaced by

$$(151) \qquad x'' - f(t)x = 0,$$

and it is not assumed that $f(t)$ is "nearly constant" for large t. The analogue of (α) is as follows:

(γ) *Let* $f(t)$ *be a continuous* (*possibly complex-valued*) *function for large* t *satisfying*

$$(152_1) \qquad \operatorname{Re} f^{\frac{1}{2}}(t) \neq 0; \qquad\qquad (152_2) \qquad \int^{\infty} |\operatorname{Re} f^{\frac{1}{2}}(t)|\, dt = \infty.$$

In addition, suppose that $f(t)$ *has a continuous derivative satisfying*

$$(153) \qquad f'(t) = o(|f(t)| \cdot |\operatorname{Re} f^{\frac{1}{2}}(t)|) \quad \text{as} \quad t \to \infty.$$

Then (151) *has a pair of solution satisfying*

$$(154)\qquad x'(t) \sim \pm f^{\frac{1}{2}}(t)x(t) \quad \text{as} \quad t \to \infty.$$

If condition (152_1) is replaced by $|\arg f(t)| < \text{const.} < \pi$ for large t, then (153) is equivalent to $f' = o(|f|^{\frac{3}{2}})$ as $t \to \infty$. When $f(t)$ is real-valued (hence positive), the condition that $f(t)$ is continuously differentiable can be omitted and (153) can be relaxed to

$$(155)\qquad \underset{0 \leqq V < \infty}{\text{l. u. b.}} \, |\log f(U+V)/f(U)|/(1 + \int_U^{U+V} f^{\frac{1}{2}}(t)dt) \to 0 \quad \text{as} \quad U \to \infty.$$

In fact, when (152) holds with $f = \operatorname{Re} f$, then (155) is necessary and sufficient for (151) to be non-oscillatory and to possess a pair of solutions satisfying (154). These assertions will be clear from the proof of (γ) and the proof of (IV) in [2], pp. 570-573.

In the proof of (γ), the condition

$$(156)\qquad \underset{0 \leqq V < \infty}{\text{l. u. b.}} \, (\int_U^{U+V} |df(t)/f(t)|)/(1 + \int_U^{U+V} |\operatorname{Re} f^{\frac{1}{2}}(t)|\, dt) \to 0 \text{ as } U \to \infty$$

will be used, rather than the more stringent condition (153).

The analogue of (β) is the following:

(δ) *If condition* (153) *in* (γ) *is replaced by*

$$(157)\qquad \int^{\infty} |f'/f|^{\gamma}\, |\operatorname{Re} f^{\frac{1}{2}}|^{\gamma-1}\, dt < \infty \text{ for some } \gamma, \text{ where } 1 \leqq \gamma \leqq 2,$$

then (151) *has a pair of solutions satisfying* (154) *and*

$$(158)\qquad x(t) \sim f^{-\frac{1}{4}}(t) \exp \pm \int^{t} f^{\frac{1}{2}}(u)\,du \quad \text{as} \quad t \to \infty.$$

In (157)-(158), as well as (154) above, $f^{\frac{1}{2}}(t)$ denotes a fixed continuous determination of the square root of f. Similarly, $f^{-\frac{1}{4}}(t)$ in (158) is a continuous square root of $1/f^{\frac{1}{2}}$.

The asymptotic formulae (158) were obtained in [10] under conditions quite different from (157). The conditions of [10] require that $f(t)$ be real-valued and either positive or negative (the latter case is eliminated from (δ) by condition (152_1) and that $f(t)$ should have a continuous second derivative satisfying (166) below.

In order to prove (γ), let $y = (y^1, y^2)$ denote the vector with the components

$$(159)\qquad y^1 = x' + f^{\frac{1}{2}}x, \qquad y^2 = x' - f^{\frac{1}{2}}x,$$

where $f^{\frac{1}{2}}$ denotes a fixed continuous determination of the square root of f. By (159),

$$x = \tfrac{1}{2} f^{-\frac{1}{2}}(y^1 - y^2), \qquad x' = \tfrac{1}{2}(y^1 + y^2). \tag{160}$$

Thus (151) is equivalent to the linear system

$$y^{1\prime} = f^{\frac{1}{2}} y^1 + f'(y^1 - y^2)/4f, \qquad y^{2\prime} = -f^{\frac{1}{2}} y^2 - f'(y^1 - y^2)/4f$$

of first order for y (where $f'/f = 2(f^{\frac{1}{2}})'/f^{\frac{1}{2}}$). Introduce the new independent variable defined by

$$s = \int^{t} g(t)\,dt, \quad \text{where} \quad g = |\operatorname{Re} f^{\frac{1}{2}}|, \qquad (ds = g\,dt). \tag{161}$$

Then $t \rightarrow \infty$ is equivalent to $s \rightarrow \infty$, by virtue of (152_2). If a dot denotes differentiation with respect to s, then the system for y becomes $\dot{y} = (J(t) + G(t))y$, where $t = t(s)$ and

$$J(t) = \begin{pmatrix} f^{\frac{1}{2}}/g & 0 \\ 0 & -f^{\frac{1}{2}}/g \end{pmatrix} \text{and } G(t) = (f'/4fg)\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Since $|\operatorname{Re}(f^{\frac{1}{2}}/g + f^{\frac{1}{2}}/g)| = 2 \neq 0$, condition (54) of (II*) is satisfied. In terms of s, condition (6) of (II*) is implied by (154) or (156). Thus there exist solutions $x = x(t)$ of (151) satisfying, as $t \rightarrow \infty$,

$$y^1 = o(y^2) \quad \text{or} \quad y^2 = o(y^1). \tag{162}$$

In view of (160), this proves assertion (154) of (γ).

[*Added in the proofs*, 12. 1. 1954. After the transformation of the differential equation (151) of second order into the binary system of first order in the formula line following (160), the assertion (γ) could also be deduced from a theorem of Z. Szmydtówna, Ann. de la Soc. Polonaise de Math., vol. 24 (1951), pp. 17-34.]

In order to prove (δ), note that (157) and (161) imply that $|G(t)|^{\gamma}$, as a function of s, is of class L. Hence Theorem (i) in Section 18 states that (151) has solutions $x = x(t)$ satisfying (154) and

$$y^1(t) \sim 2 \exp \int^{t} (f^{\frac{1}{2}}(u) + f'(u)/4f(u))\,du,$$

$$y^2(t) \sim 2 \exp \int^{t} (-f^{\frac{1}{2}}(u) + f'(u)/4f(u))\,du.$$

Since the latter two formulae can be written as

$$y^1(t) \sim 2f^{\frac{1}{4}}(t) \exp \int^t f^{\frac{1}{2}}(u)\,du, \qquad y^2(t) \sim 2f^{\frac{1}{4}}(t) \exp \int^t -f^{\frac{1}{2}}(u)\,du,$$

assertion (δ) follows from (159) and (161).

Another theorem leading to asymptotic formulae for the solutions of (151) is the following:

(ϵ) *Let* $f(t)$ *satisfy the conditions of* (γ). *In addition, let* $f'/f^{\frac{3}{2}}$ *be of bounded variation for large* t, *that is, let*

$$\int^\infty |\,d(f'/f^{\frac{3}{2}})\,| < \infty. \tag{163}$$

Then (151) *has a pair of solutions satisfying* (154) *and*

$$x(t) \sim f^{-\frac{1}{4}}(t) \exp \pm \int^t f^{\frac{1}{2}}(1 + f'^2/16f^3)^{\frac{1}{2}}dt \quad \text{as} \quad t \to \infty. \tag{164}$$

In view of (153), $f'^2/f^3 \to 0$ as $t \to \infty$. The square root in $(1 + f'^2/16f^3)^{\frac{1}{2}}$ can be chosen so that $(1 + f'^2/16f^3)^{\frac{1}{2}} \to 1$ as $t \to \infty$; the determinations of $f^{\frac{1}{2}}$ and $f^{-\frac{1}{4}}$ are chosen as in (158).

If $f(t)$ has a continuous second derivative, condition (163) means that

$$\int^\infty |\,f''/f^{\frac{3}{2}} - 3f'^2/2f^{\frac{5}{2}}\,|\,dt < \infty. \tag{165}$$

The condition in [10] for the validity of (158) (rather than (164)) is

$$\int^\infty |\,f''/f^{\frac{3}{2}} - 5f'^2/4f^{\frac{5}{2}}\,|\,dt < \infty \tag{166}$$

(when $f(t) \neq 0$ is real-valued).

In order to prove (ϵ), note that (γ) is applicable, hence (151) has solutions satisfying (154). Let $x = x(t)$ be a solution of (151) which satisfies

$$v(t) \to 1 \quad \text{as} \quad t \to \infty, \tag{167}$$

where

$$v = x'/f^{\frac{1}{2}}x. \tag{168}$$

It is readily verified that (168) satisfies the Riccati equation

$$v' + v^2 f^{\frac{1}{2}} + \tfrac{1}{2}f'f^{-1}v - f^{\frac{1}{2}} = 0,$$

by virtue of (151). This can be written as

$$v' + f^{\frac{1}{2}}\{(v + f'/4f^{\frac{3}{2}})^2 - (1 + f'^2/16f^3)\} = 0,$$

or as

$$(169)\qquad f^{\frac{1}{2}}\{(v+f'/4f^{\frac{3}{2}})-(1+f'^{2}/16f^{3})^{\frac{1}{2}}\}=-v'/(v+a(t)),$$

where

$$(170)\qquad a(t)=f'/4f^{\frac{3}{2}}+(1+f'^{2}/16f^{3})^{\frac{1}{2}}.$$

Note that $a(t)\rightarrow 1$ as $t\rightarrow\infty$, by the above normalizations of the square roots; so that $v(t)+a(t)\rightarrow 2$ as $t\rightarrow\infty$, by (167).

Let v on the left of (169) be written as $x'/f^{\frac{1}{2}}x$ (cf. (168)). Then a quadrature shows that

$$x(t)f^{\frac{1}{4}}(t)/\exp\int^{t} f^{\frac{1}{2}}(u)(1+f'^{2}(u)/16f^{3}(u))^{\frac{1}{2}}du$$

is equal to

$$(a(t)+v(t))^{-1}\exp\int^{t}(v(u)+a(u))^{-1}da(u).$$

This is clear if $v'dt$ is written as $d(v+a)-da$. The expression in the last formula line tends to a limit, as $t\rightarrow\infty$, since $(a(t)+v(t))^{-1}$ does and since $a(u)$ is of bounded variation. Thus if const. $x(t)$, for a suitable choice of const., is renamed $x(t)$, it is clear that (164) holds when $+$ is written in place of $\pm$. Since the choice of the square root $f^{\frac{1}{2}}$ in (168) was arbitrary, (ϵ) is proved.

25. An elliptic analogue. Assertions (β) or (δ) have no analogues when λ is a negative constant or $f(t)$ is negative for large t, respectively. As to (β), cf. the example in [11], p. 269. On the other hand, (ϵ) does have an analogue in the "elliptic" case.

(η) *Let* $f(t)$ *be a non-vanishing (possibly complex-valued) function for large* t *having a continuous derivative satisfying* (163) *and*

$$(171)\qquad \lim_{t\rightarrow\infty} f'^{2}/16f^{3}\neq -1.$$

Suppose further that

$$(172)\qquad \exp\pm i\int^{t} f^{\frac{1}{2}}(1+f'^{2}/16f^{3})^{\frac{1}{2}}du=O(1)\quad\text{as}\quad t\rightarrow\infty.$$

Then the differential equation

$$(173)\qquad x''+f(t)x=0$$

has a pair of solutions satisfying, as $t\rightarrow\infty$,

$$(174)\qquad x\sim f^{-\frac{1}{4}}(t)\exp\pm i\int^{t} f^{\frac{1}{2}}(1+f'^{2}/16f^{3})^{\frac{1}{2}}du$$

and

$$x' \sim (-\alpha^{\frac{1}{2}} \pm i(1+\alpha)^{\frac{1}{2}})f^{\frac{1}{2}}x, \tag{175}$$

where α *is the limit in* (171).

Conditions (171)-(172) are trivially satisfied if $f(t)$ is (real and) positive for large t. As in the above theorems, all powers of $f(t)$ which occur can be assumed to be integral (positive or negative) powers of a fixed fourth root $f^{\frac{1}{4}}(t)$; in such a normalization, $\alpha^{\frac{1}{2}}$ in (175) is the limit, as $t \rightarrow \infty$, of $(f^{\frac{1}{4}})'/f^{\frac{3}{4}}$.

The proof of (η) will depend on the simplest case, $J=0$, of Theorem (**) in Part II and will be like the procedure used in [11], pp. 262-263, for similar purposes.

In the proof of (η), there is no actual loss of generality in the (local) assumption that $f(t)$ has a continuous second derivative; so that (163) is equivalent to (165). The variation of constants

$$z = xf^{\frac{1}{4}} \tag{176}$$

transforms (173) into

$$(f^{-\frac{1}{2}}z')' + (f^{\frac{1}{2}}(t) + a(t))z = 0, \tag{177}$$

where

$$a(t) = f''/4f^{\frac{3}{2}} - 5f'^2/16f^{\frac{5}{2}}. \tag{178}$$

If $y = (y^1, y^2)$, where

$$y^1 = z, \qquad y^2 = f^{-\frac{1}{2}}z', \tag{179}$$

then (177) is equivalent to the system

$$y' = A(t)y, \quad \text{where} \quad A(t) = \begin{pmatrix} 0 & f^{\frac{1}{2}} \\ -f^{\frac{1}{2}} - a & 0 \end{pmatrix}. \tag{180}$$

If z denotes either of the functions on the left of (172) and if

$$(\cdot\ \cdot\ \cdot) = (1 + f'^2/16f^3), \tag{181}$$

then $z' = \pm if^{\frac{1}{2}}(\cdot\ \cdot\ \cdot)^{\frac{1}{2}}z$, and so $(f^{-\frac{1}{2}}z')' = -(\cdot\ \cdot\ \cdot)f^{\frac{1}{2}}z \pm i\{(\cdot\ \cdot\ \cdot)^{\frac{1}{2}}\}'z$. If $\pm iz$ in the last term is replaced by $z'f^{-\frac{1}{2}}(\cdot\ \cdot\ \cdot)^{-\frac{1}{2}}$, it is seen that the functions on the left of (172) are solutions of the differential equation

$$(f^{-\frac{1}{2}}z') - \tfrac{1}{2}f^{-\frac{1}{2}}(\cdot\ \cdot\ \cdot)^{-1}(\cdot\ \cdot\ \cdot)'z' + (\cdot\ \cdot\ \cdot)f^{\frac{1}{2}}z = 0. \tag{182}$$

Let $w = (w^1, w^2)$, where

$$w^1 = z, \qquad w^2 = f^{-\frac{1}{2}}z'; \tag{183}$$

so that (182) is equivalent to the system

$$(184)\qquad w' = B(t)w,\text{ where } B(t) = \begin{pmatrix} 0 & f^{\frac{1}{2}}(t) \\ -(\cdot\cdot\cdot)f^{\frac{1}{2}} & \frac{1}{2}(\cdot\cdot\cdot)^{-1}(\cdot\cdot\cdot)' \end{pmatrix}.$$

In view of the derivation of the differential equation (182), a fundamental matrix for (182) is

$$(185)\qquad W(t) = \begin{pmatrix} w_1{}^1(t) & w_2{}^1 \\ -i(\cdot\cdot\cdot)^{\frac{1}{2}}w_1{}^1 & i(\cdot\cdot\cdot)w_2{}^1 \end{pmatrix},$$

where

$$(186)\qquad w_j{}^1 = \exp(-1)^j i \int^t f^{\frac{1}{2}}(1 + f'^2/16f^3)^{\frac{1}{2}}du, \qquad j = 1, 2,$$

are the functions in (172).

The variation of constants

$$(187)\qquad y = W(t)v$$

transforms (180) into

$$(188)\qquad v' = W^{-1}(A - B)Wv.$$

In view of (178), (180), (181) and (184),

$$A - B = \begin{pmatrix} 0 & 0 \\ 3f'^2/8f^{\frac{5}{2}} - f''/4f^{\frac{3}{2}} & -\frac{1}{2}(\cdot\cdot\cdot)^{-1}(\cdot\cdot\cdot)' \end{pmatrix}.$$

The assumption (163) implies that the elements of $A - B$ are of class L (cf. (165)), while the condition (172) implies that $W(t) = O(1)$ as $t \to \infty$. Since $\det W(t) = 2i(\cdot\cdot\cdot)$, it follows that $(\det W(t))^{-1}$, and therefore $W^{-1}(t)$, is $O(1)$, as $t \to \infty$. Thus the coefficients of the system (188) are of class L for large t. Hence, the simplest case, $J = 0$, of (**) states that (188) has a solution satisfying, as $t \to \infty$,

$$(189)\qquad v^1(t) \to 1 \quad\text{and}\quad v^2(t) \to 0,$$

and another solution satisfying, as $t \to \infty$,

$$(190)\qquad v^1(t) \to 0 \quad\text{and}\quad v^2(t) \to 1.$$

The solution $y = y(t)$ of (180) corresponding to (189), by virtue of (187), satisfies

$$y^1(t) \sim w_1{}^1(t), \qquad y^2(t) \sim -i(\cdot\cdot\cdot)^{\frac{1}{2}}w_1{}^1$$

as $t \to \infty$. In view of (176) and (179), the corresponding solution $x = x(t)$ of (173) satisfies

$$f^{\frac{1}{4}}x \sim w_1{}^1, \qquad f^{-\frac{1}{2}}(xf^{\frac{1}{4}})' \sim -i(\cdot\cdot\cdot)^{\frac{1}{2}}w_1{}^1$$

as $t \to \infty$. The second of these relations reduces, by virtue of the first, to

$$x' = (1 + o(1))\{-f'/4f^{\frac{3}{2}} - i(\cdot\ \cdot\ \cdot)^{\frac{1}{2}}\}f^{\frac{1}{2}}x.$$

The last two formula lines prove the case — of ± in (174), (175). The other case results from (190). This proves (η).

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] O. Dunkel, "Regular singular points of a system of homogeneous linear differential equations of the first order," *Proceedings of the American Academy of Arts and Sciences*, vol. 38 (1902-03), pp. 341-370.

[2] P. Hartman, "Unrestricted solution fields of almost separable differential equations," *Transactions of the American Mathematical Society*, vol. 63 (1948), pp. 560-580.

[3] ———, "On non-oscillatory linear differential equations of second order," *American Journal of Mathematics*, vol. 74 (1952), pp. 389-400.

[4] ——— and A. Wintner, "The asymptotic arcus variation of solutions of real linear differential equations of second order," *ibid.*, vol. 70 (1948), pp. 1-10.

[5] ——— and A. Wintner, "On non-linear differential equations of first order," *ibid.*, vol. 72 (1950), pp. 347-358.

[6] ——— and A. Wintner, "On non-oscillatory differential equations," *ibid.*, vol. 75 (1953), pp. 717-730.

[7] F. Lettenmeyer, "Ueber das asymptotische Verhalten der Lösungen von Differentialgleichungen und Differentialgleichungssystemen," *Sitzungsberichte München*, 1929, pp. 201-252.

[8] O. Perron, "Über Stabilität und asymptotisches Verhalten der Integrale von Differentialgleichungssystemen," *Mathematische Zeitschrift*, vol. 29 (1929), pp. 129-160.

[9] ———, "Die Ordnungszahlen linearer Differentialgleichungssysteme," *ibid.*, vol. 31 (1930), pp. 748-766.

[10] A. Wintner, "On the normalization of characteristic differentials in continuous spectra," *Physical Review*, vol. 72 (1947), pp. 516-517.

[11] ———, "Asymptotic integrations of the adiabatic linear oscillator," *American Journal of Mathematics*, vol. 69 (1947), pp. 251-272.

QUADRATIC FORMS OVER LOCAL FIELDS.*

By O. T. O'Meara.

Introduction. Even after Minkowski had *solved* the problem of equivalence of quadratic forms the matter was by no means *settled*, for the invariants found in [17] are intricate both in their derivation and application. The difficulties presented (even in the theory of rational equivalence) were largely due to the use of integral forms and the problem was simplified by Hasse [5] who rid it of its integral considerations, solved the problem of fractional equivalence in the p-adic numbers and tied them together to form a complete system of invariants under rational equivalence in the large; he related these invariants to those of Minkowski. In the subsequent paper [6], Hasse extended his results to quadratic forms over algebraic number fields and their local completions. These results were in turn extended by Witt's using the theory of algebras in [20]. Kaplansky showed in [13] that Witt's invariants apply to fields F which are not formally real and for which $F^*/(F^*)^2$ has order 4.

The problem of tackling the original theory of integral equivalence remained, but use could now be made of the results on fractional equivalence proved independently of it. Siegel [18] proved a conjecture of Minkowski that two rational integral forms are in the same genus if and only if they are semi-equivalent; Jones [7] proved that the Witt cancellation law [20] holds under integral equivalence over the p-adic numbers (p odd) and stated that the result also held for arbitrary local fields in which $|2|=1$; Durfee [3, 4] gave a complete set of invariants and a simple proof of the Witt cancellation law for integral forms in these fields; Jones [8] showed how to determine when two forms over the 2-adic integers are equivalent and this was expressed by Pall [15] in the form of invariants; most of these results can be found in coordinated form (but over the rational p-adic numbers only) in [9].

Eichler employed the more geometrical notion of lattices to discuss the local theory in [11, Chapter II]: of particular interest to us will be his results on the equivalence of maximal forms. Eichler [12] and Kneser [14] proved the Witt cancellation law for positive definite forms over the rational integers.

---

* Received August 26, 1953; revised June 8, 1954 and September 15, 1954.

In the present paper we are concerned with integral equivalence over local fields and the previous remarks show that the only local cases requiring real investigation are those in which $|2| < 1$. We prove a weak form of cancellation law for these fields: that this cannot be fully generalized is clear from the counter-example in [3]. Certain useful type invariants are given in the general case, but these do not form a complete set under integral equivalence. However a characterization is given over local fields in which 2 is a prime; this is facilitated by the use of lattices but the real key seems to lie in the perfectness of the residue class field. The ramified case is much more complicated than this and simple examples show that we cannot expect the same answers as in the unramified theory; the numbers represented by the form seem to play an important role; the perfectness of the residue class field pulls us through again and we can characterize forms of unit determinant, but the problem remains of finding a complete set of invariants for forms of arbitrary determinant over a general local field.

1. Preliminary definitions. In the present paper we consider quadratic forms over local fields, that is to say over fields which are complete under a discrete non-Archimedean valuation and have a perfect residue class field [1]. F will denote the local field, $\mathfrak{o}$ the ring of integers in F and π will be a fixed prime element in F. We define $e = \operatorname{ord}(2)$ as the ramification index of F. We assume in addition that F does not have characteristic 2. To avoid the use of an excessive number of multipliers we introduce the following notation: a specific element $a\pi^t$ will sometimes be written as $[\pi^t]$ if $|a| = 1$ or as $\{\pi^t\}$ if $a \varepsilon \mathfrak{o}$. Thus, if $\alpha = [\pi^t]$, $\beta = [\pi^t]$ and $\gamma = \{\pi^t\}$, α need not be equal to β but we must have $|\gamma| \leqq |\alpha| = |\beta|$.

If R is an n-dimensional vector space over F, we call an $\mathfrak{o}$-module L in R a lattice if there is a basis for R in which L can be expressed as $L = \mathfrak{o}\xi_1 + \cdots + \mathfrak{o}\xi_n$ [see 11, p. 47]. We call (ξ) a minimal basis for L and note that L may very well have many minimal bases.

Now let l be a quadratic form and $\Lambda = (l_{ij})$ its corresponding matrix. By fixing a basis (ξ) for R we can make R into a metric space V by defining $\xi_i \cdot \xi_j = l_{ij} = l_{ji} = \xi_j \cdot \xi_i$, [20, p. 32]. The form k is fractionally equivalent to l if and only if there is another basis (η) for V in which $k_{ij} = \eta_i \cdot \eta_j$. We then write $l \sim k$. Conversely we can associate a quadratic form with a given metric space V by fixing a basis (ξ) for V and defining $l_{ij} = \xi_i \cdot \xi_j$.

Now let $l \sim k$ have respective bases (ξ) and (η) in V and consider the lattices $L = \mathfrak{o}\xi_1 + \cdots + \mathfrak{o}\xi_n$ and $K = \mathfrak{o}\eta_1 + \cdots + \mathfrak{o}\eta_n$. Then L and K are integrally equivalent (written $L \cong K$) if and only if there are minimal bases (ξ') and (η') for L and K respectively such that $\xi_i' \cdot \xi_j' = \eta_i' \cdot \eta_j'$, and this in turn is equivalent to saying that there is an isometry $V \rightarrow V$ which carries L onto K. If L and K are isometric we write $L \cong K$, $l \cong k$ and $\Lambda \cong \mathrm{K}$. We denote the fact that l is associated with L by writing $l \cong L \cong \Lambda$. We call $\det(l_{ij})$ the determinant of L and write it as $d(L)$; this is unique but for a factor which is the square of a unit. We assume that all forms are non-degenerate: $\det(l_{ij})$ unequal to zero.

L is the orthogonal sum of sublattices L_1 and L_2, written $L = L_1 \oplus L_2$, if $L_1 \cap L_2 = 0$, L_1 is orthogonal to L_2, and L_1 and L_2 together span L. It is easy to see that if $L = L_1 \oplus K_2$ then $K_2 = L_2$.

The ideal generated by X^2 as X runs through L will be called the norm of L and written $N(L)$. L is said to be integral if $N(L) \subseteq \mathfrak{o}$: if L is integral, then $X \cdot Y \in (\frac{1}{2}) \cdot \mathfrak{o}$ for all X and Y in L. If in addition $X \cdot Y \in \mathfrak{o}$ for all X and Y in L, we say that L is totally integral.

I. Lattices Over Local Fields.

In this chapter we discuss type invariants, lattices of unit determinant and the cancellation law. In general the fields considered will be local fields but it will be necessary to indicate certain modifications for different values of e.

2. Canonical representation of a lattice. Our aim is to show that every lattice $L = \mathfrak{o}\xi_1 + \cdots + \mathfrak{o}\xi_n$ can be expressed in an almost orthogonal canonical basis.

Definition 2.1. $\xi_1, \cdots, \xi_n$ is a canonical basis for L if

(1) when $e = 0$: $\xi_i \cdot \xi_j = 0$ for all $i \neq j$; and $|\xi_i^2| \geqq |\xi_{i+1}^2|$.

(2) when $e \geqq 1$: (a) $\xi_i \cdot \xi_j \neq 0$ and $i \neq j$ happens in at most one of the cases $j = i - 1$ or $j = i + 1$; and $\xi_i \cdot \xi_j$ is then an exact power of π; (b) $\max |\xi_i \cdot Z| \geqq \max |\xi_{i+1} \cdot Z|$, $Z \in L$, $1 \leqq i \leqq n - 1$; (c) if $\max |\xi_i \cdot Z| = |\xi_j^2| = \max |\xi_j \cdot Z|$, $Z \in L$, holds for a particular value of i and a particular value of j (which may be equal to i), then that ξ_i is orthogonal to all ξ_k, $k \neq i$, and so $|\xi_i^2| = \max |\xi_i \cdot Z| = |\xi_j^2|$. Observe that (a), (b) and (c) imply that if $\xi_i \cdot \xi_{i+1} \neq 0$ then $|\xi_i^2| < |\xi_i \cdot \xi_{i+1}|$ and $|\xi_{i+1}^2| < |\xi_i \cdot \xi_{i+1}|$.

We now prove some results for the lattice L expressed in any basis (ξ).

LEMMA 2.1. *If* $|\xi_1^2| = \max |\xi_1 \cdot \xi_i|$, *then there is a sublattice* L_1 *of* L *such that* $L = \mathfrak{o}\xi_1 \oplus L_1$.

Proof. Replace ξ_i by $\Xi_i = (\xi_i - (\xi_1 \cdot \xi_i)/\xi_1^2 \cdot \xi_1)$ for all $1 < i \leqq n$. Since $\xi_1 \cdot \xi_i/\xi_1^2$ is an integer it follows that $L = \mathfrak{o}\xi_1 + \mathfrak{o}\Xi_2 + \cdots + \mathfrak{o}\Xi_n$. The remark $\xi_1 \cdot \Xi_i = 0$ completes the proof.

LEMMA 2.2. *If the equations* $\xi_1 \cdot \xi_i + \alpha_i\xi_1^2 + \beta_i(\xi_1 \cdot \xi_2) = 0$ *and* $\xi_2 \cdot \xi_i + \alpha_i(\xi_1 \cdot \xi_2) + \beta_i\xi_2^2 = 0$ *have a simultaneous integral solution* (α_i, β_i) *for all* $2 < i \leqq n$, *then there is a sublattice* L_1 *of* L *such that* $L = (\mathfrak{o}\xi_1 + \mathfrak{o}\xi_2) \oplus L_1$.

Proof. Replace each ξ_i, $2 < i \leqq n$, by $\Xi_i = \xi_i + \alpha_i\xi_1 + \beta_i\xi_2$, and write $L = (\mathfrak{o}\xi_1 + \mathfrak{o}\xi_2) \oplus (\mathfrak{o}\Xi_3 + \cdots + \mathfrak{o}\Xi_n)$. Q. E. D.

LEMMA 2.3. *Let* $e > 0$. *If* L *is a totally integral lattice of unit determinant and if* $N(L) = \mathfrak{o}$, *then* L *has an orthogonal basis.*

Proof. Induction to n. Since $N(L) = \mathfrak{o}$ there must be a ξ_i (call it ξ_1) such that ξ_1^2 is a unit. Use Lemma 2.1 and write $L = \mathfrak{o}\xi_1 \oplus L_1$. If $N(L_1) = \mathfrak{o}$, we are through. If $N(L_1) \subseteq (\pi) \cdot \mathfrak{o}$, let ξ_2, ξ_3 be two basis elements of L_1 such that $\xi_2 \cdot \xi_3$ is a unit (two such vectors must exist since $d(L_1)$ is a unit); L_1 must necessarily be of the form $(\mathfrak{o}\xi_2 + \mathfrak{o}\xi_3) \oplus L_2$ since $|\xi_2 \cdot \xi_3| = 1$, $\xi_2^2 \,\varepsilon\, (\pi) \cdot \mathfrak{o}$, $\xi_3^2 \,\varepsilon\, (\pi) \cdot \mathfrak{o}$ allows integral solutions for the equations of Lemma 2.2. Then

$$\mathfrak{o}\xi_1 + \mathfrak{o}\xi_2 + \mathfrak{o}\xi_3 = \mathfrak{o}(\xi_2 + \xi_1) + \mathfrak{o}\xi_3 + \mathfrak{o}\xi_1 = (\mathfrak{o}(\xi_2 + \xi_1) + \mathfrak{o}\xi_3) \oplus \mathfrak{o}\eta,$$

again in virtue of Lemma 2.2. Hence $L = \mathfrak{o}\eta \oplus L_1^*$ with $(\xi_2 + \xi_1) \,\varepsilon\, L_1^*$; that is, $N(L_1^*) = \mathfrak{o}$. Induction completes the proof.

THEOREM 2.4. *L has a canonical basis.*

Proof. Induction to n. There is no loss of generality if we assume that L is totally integral and contains an X and a Y for which $X \cdot Y$ is a unit. If $e = 0$ this implies that there is a basis in which ξ_1^2 (say) is a unit.

Case 1. $N(L) = \mathfrak{o}$. By the previous remark this includes the case $e = 0$ and there is a ξ_1 such that $|\xi_1^2| = 1$. If $e \geqq 1$ there must still be a ξ_1 (say) such that $|\xi_1^2| = 1$ else $N(L)$ would not be $\mathfrak{o}$. Then Lemma 2.1 entitles us to write $L = \mathfrak{o}\xi_1 \oplus L_1$. The inductive assumption allows us to express L_1 in a canonical basis in such a way that $L_1 = L_1^* \oplus L_1^{**}$ where L_1^* is of unit determinant and $|X \cdot Y| < 1$ for all X and Y in L_1^{**}. By the induction when $e = 0$, and by Lemma 2.3 when $e \geqq 1$, $\mathfrak{o}\xi_1 \oplus L_1^*$ must have an orthogonal basis $\eta_1, \cdots, \eta_i$ with the property $|\eta_j^2| = 1$ for all $1 \leqq j \leqq i$; and L_1^{**} has a canonical basis $\eta_{i+1}, \cdots, \eta_n$. Then

$$L = (\mathfrak{o}\eta_1 + \cdots + \mathfrak{o}\eta_i) \oplus (\mathfrak{o}\eta_{i+1} + \cdots + \mathfrak{o}\eta_n)$$

is the basis required for L.

Case 2. $N(L) \subseteq (\pi)\cdot\mathfrak{o}$. Let $|\xi_1 \cdot \xi_2| = 1$ (say). Then $\xi_1{}^2 \in (\pi)\cdot\mathfrak{o}$ and we can therefore apply Lemma 2. 2 to express $L = (\mathfrak{o}\xi_1 + \mathfrak{o}\xi_2) \oplus L_1$, replacing ξ_1 by $\xi_1/(\xi_1\cdot\xi_2)$ if necessary in order to ensure that $\xi_1\cdot\xi_2 = \pi^0$. Induction yields the result and the proof is therefore complete.

Of course L will have many canonical bases, but we aim to show that it can be characterized in terms of invariants attached to an arbitrary canonical basis provided that $e \leqq 1$. Lemmas 2. 1 and 2. 2 are of practical importance; they are proved in [11, Satz 9. 2] but the canonical form established there is not the same as ours and need not satisfy definition 2. 1, (b) and (c).

Having expressed L in a canonical basis we can write

$$(2.1)\quad L = L_1(\pi^{s(1)}) \oplus L_2(\pi^{s(2)}) \oplus \cdots \oplus L_r(\pi^{s(r)}) \text{ with } s(i) < s(i+1)$$

where $L_j(\pi^{s(j)})$ is the sublattice spanned by these ξ_i for which $\max|\xi_i\cdot Z| = |\pi^{s(j)}|$; then $L_j \cong \pi^{s(j)}\cdot l_j$ where l_j is totally integral, of unit determinant and either of the form $\sum[1]$ or $\sum\begin{pmatrix}\{\pi\} & 1\\ 1 & \{\pi\}\end{pmatrix}$; and $s(1) < s(2) \cdots < s(r)$.

Definition 2. 2. $L_m(\pi^{s(m)})$ is a proper lattice if $N(L_m) = (\pi^{s(m)})\cdot\mathfrak{o}$; otherwise improper.

A glance at the determinant of L_m together with Theorem 2. 4 shows that L_m has an orthogonal basis if and only if it is proper; if improper, L_m is the orthogonal sum of indecomposable lattices of dimension precisely 2. Improper lattices do not occur when $e = 0$.

3. Invariance of type. In order to be able to attach certain well-defined quantities to L we first prove the following theorem which we have already shown to hold for lattices of unit determinant. Compare the weaker result [3, Lemma 2] and the p-adic result of Jones [9, Theorem 35]; observe that Theorem 3. 1 strengthens [3, Lemma 4].

THEOREM 3. 1. *Let*

$$L_1(\pi^{s(1)}) \oplus \cdots \oplus L_r(\pi^{s(r)}) \quad \textit{and} \quad K_1(\pi^{t(1)}) \oplus \cdots \oplus K_q(\pi^{t(q)})$$

be two canonical representations of L. *Then* $s(i) = t(i)$, $\dim L_i = \dim K_i$ *and* $N(L_i) \equiv N(K_i) \pmod{\pi^{s(i)+1}}$ *for all* i.

Proof. By considering $N(L)$ we see that $s(1)$ must be equal to $t(1)$. We suffer no loss in generality in assuming that $s(1) = 0 = t(1)$ and consequently that $N(L) \subseteq \mathfrak{o}$. Let (ξ) and (η) be respective canonical bases for L. First we show that $\dim L_i = \dim K_i$ and $s(i) = t(i)$ all i: by definition of L_i and K_i this is tantamount to proving max $| \xi_j \cdot Z | =$ max $| \eta_j \cdot Z |$ all j, $Z \varepsilon L$. If not, let ξ_d be the first point of discrepancy: max $| \xi_d \cdot Z | <$ max $| \eta_d \cdot Z | = | \pi^\delta |$, say. Then expressing (η_j) in terms of the (ξ_j) we see that

$$\eta_1 = a_{11}\xi_1 + \cdots + a_{1d-1}\xi_{d-1} + N_1$$

$$\cdots \qquad \cdots \qquad \cdots$$

$$\eta_d = a_{d1}\xi_1 + \cdots + a_{dd-1}\xi_{d-1} + N_d$$

where max $| N_i \cdot Z | < | \pi^\delta |$ and hence by adding vertically we see that there must exist integers λ_γ with at least one λ_γ a unit such that $\sum^d \lambda_\gamma \cdot \eta_\gamma = \sum^d \lambda_\gamma \cdot N_\gamma$ Now max $| N_\gamma \cdot Z | < | \pi^\delta |$ and max $| \eta_\gamma \cdot Z | \geqq | \pi^\delta |$ provided that $\gamma \leqq d$. Since one of the λ_γ is a unit this establishes a contradiction and the first contention is proved.

Suppose now that i is an integer for which $L_i(\pi^t)$ and $K_i(\pi^t)$ are of different *type*: let $K_i = o\eta_\mu + \cdots$ be proper. Then there is an $N \varepsilon L$ with the property that $| N^2 | = | \pi^t |$ and $| N \cdot Z | \leqq | \pi^t |$ for all $Z \varepsilon L$. We aim to exhibit an M orthogonal to $\xi_1, \cdots, \xi_{\mu-1}$ such that $| M^2 | = | \pi^t |$. Put $M = (N - \lambda_1 \cdot \xi_1 \cdots - \lambda_{\mu-1}\xi_{\mu-1})$ where λ_j are integers defined for $j < \mu$ in the following way: (1) if $\xi_i \cdot \xi_k = 0$ all $k \neq j$ write $\lambda_j = (N \cdot \xi_j)/\xi_j^2$. Then $| \lambda_j | \leqq | \pi^t |/| \xi_j^2 |$ and λ_j is therefore an integer since $j < \mu$. In addition $| (\lambda_j \cdot \xi_j)^2 | \leqq | \pi^{2t} |/| \xi_j^2 | < | \pi^t |$ since $j < \mu$; and similarly $| 2\lambda_j \cdot (N \cdot \xi_j) | \leqq | \pi^{2t} |/| \xi_j^2 | < | \pi^t |$. (2) if on the other hand ξ_j is not orthogonal to ξ_{j-1} (say), choose λ_j and λ_{j-1} as the *integral* solutions of the equations

$$N \cdot \xi_j = \lambda_j \cdot \xi_j^2 + \lambda_{j-1} \cdot (\xi_j \cdot \xi_{j-1})$$

$$N \cdot \xi_{j-1} = \lambda_j \cdot (\xi_j \cdot \xi_{j-1}) + \lambda_{j-1} \cdot \xi_{j-1}^2.$$

Then $| \xi_j^2 \cdot \xi_{j-1}^2 - (\xi_j \cdot \xi_{j-1})^2 | = | (\xi_j \cdot \xi_{j-1})^2 |$ since $| \xi_j^2 | < | \xi_j \cdot \xi_{j-1} |$ and $| \xi_{j-1}^2 | < | \xi_j \cdot \xi_{j-1} |$, (ξ) being a canonical basis; direct computation then shows that $| \lambda_j | \leqq | \pi^t |/| \xi_j \cdot \xi_{j-1} |$ and $| \lambda_{j-1} | \leqq | \pi^t |/| \xi_j \cdot \xi_{j-1} |$, so that λ_j and λ_{j-1} are certainly integers. Multiplication shows that $| (\lambda_j \cdot \xi_j + \lambda_{j-1} \cdot \xi_{j-1})^2 | < | \pi^t |$ and $| 2N \cdot (\lambda_j \cdot \xi_j + \lambda_{j-1} \cdot \xi_{j-1}) | < | \pi^t |$.

The properties proved in (1) and (2) show that $M^2 \equiv N^2 \pmod{\pi^{t+1}}$ and hence $| M^2 | = | \pi^t |$. In addition the λ_j in both (1) and (2) were chosen in such a way that $M \cdot \xi_j = 0$ all $j < \mu$. Hence $L_i(\pi^t)$ must represent a number of ordinal t. Consequently $L_i(\pi^t)$ must be proper. This contradiction completes the proof.

Definition 3.1. Let L and K be two lattices having canonical representations

$$L_1(\pi^{s(1)}) \oplus \cdots \oplus L_r(\pi^{s(r)}) \text{ and } K_1(\pi^{t(1)}) \oplus \cdots \oplus K_r(\pi^{t(r)})$$

respectively, with $s(1) < \cdots < s(r)$ and $t(1) < \cdots < t(r)$. We say that L and K are of the same type if $s(i) = t(i)$, $\dim L_i = \dim K_i$ and L_i and K_i are proper or improper together, where $1 \leqq i \leqq r$. We call $|\pi^{s(r)}| = |\pi^{t(r)}|$ the least divisor of the type and write $s(r) = m(L)$; we call $|\pi^{s(1)}| = |\pi^{t(1)}|$ the greatest divisor of the type and write $s(1) = M(L)$.

4. Lattices of unit determinant. We are now in a position to characterize lattices of unit determinant over fields in which 2 is not ramified. But first let us state some well-known lemmas in a form more suited to our purposes.

LEMMA 4. 1. *Let $e > 0$. If ϵ is a unit in F and if X is a vector in L, then the equation* $(\alpha \cdot X)^2 \equiv \epsilon X^2 \pmod{\pi X^2}$ *always has a solution α.*

Proof. The required α is any solution of $\alpha^2 \equiv \epsilon \pmod{\pi}$; this always exists in virtue of the fact that the residue class field is perfect, [19].

LEMMA 4. 2. *If* $X^2 \equiv \epsilon X^2 \pmod{4\pi \cdot X^2}$, *then there is an $\alpha \in F$ such that* $(\alpha \cdot X)^2 = \epsilon \cdot X^2$.

Proof. $\epsilon \equiv 1 \pmod{4\pi}$ implies that $\alpha^2 = \epsilon$ has a solution α in F by Hensel's lemma [1].

LEMMA 4. 3. *Let $L = \mathfrak{o}\xi_1 + \cdots + \mathfrak{o}\xi_n$ and let (η_i) be another basis for V. Then there is minimal basis (ζ_i) for L with the property $\zeta_i \in F \cdot \eta_1 + \cdots + F \cdot \eta_i$.*

Proof. By induction to n. Write $\xi_i = a_{i1} \cdot \eta_1 + \cdots + a_{in} \cdot \eta_n$ and permute the (ξ_i) until $|a_{nn}| = \max |a_{\lambda n}|$. Then L can be written as $\mathfrak{o}\Xi_1 + \cdots + \mathfrak{o}\Xi_{n-1} + \mathfrak{o}\xi_n$ where $\Xi_i = (\xi_i - a_{in}/a_{nn} \cdot \xi_n) \in F \cdot \eta_1 + \cdots + F \cdot \eta_{n-1}$. Induction completes the proof. [See [2], Theorems 1 and 2.]

Now we give the characterization mentioned above.

THEOREM 4. 4. *Let L and K be totally integral lattices of unit determinant in the same metric space V. If $N(L) = (2) \cdot \mathfrak{o} = N(K)$ then $L \cong K$. If $e = 1$ and $N(L) = N(K)$ then $L \cong K$.*

Proof. If $N(L)$ and $N(K)$ are both equal to $(2) \cdot \mathfrak{o}$ then L and K are maximal lattices [11] of norm $(2) \cdot \mathfrak{o}$ and hence $L \cong K$ by [11, Sätz 9. 6]. This proves the first part of the theorem and from now on we can take $e = 1$. There are two possibilities: (1) if L is proper then $N(L) = \mathfrak{o} = N(K)$ and K is therefore proper too; (2) if L is improper then L contains a vector of norm $(2) \cdot \mathfrak{o}$: for if $\xi_1, \xi_2, \cdots$ denotes a canonical basis for L then $|\xi_1{}^2|$ or

$|\xi_2^2|$ or $|(\xi_1+\xi_2)^2|$ must be equal to $|2|$; hence $N(L)=(2)\cdot\mathfrak{o}=N(K)$ and this possibility has already been proved. The problem therefore reduces to the case (1) where L and K are both proper: let $L=\mathfrak{o}\xi_1+\cdots+\mathfrak{o}\xi_n$ and $K=\mathfrak{o}\zeta_1+\cdots+\mathfrak{o}\zeta_n$ and proceed by induction to n.

If we assume that either L or K does not represent $\{\pi^2\}$ non-trivially (let us say L) then we can show that L is precisely K as follows: since L and K are in the same space we have $\zeta_j=a_{1j}\cdot\xi_1+\cdots+a_{nj}\cdot\xi_n$ and suppose that $|a_{ij}|=\max|a_{\lambda j}|$ is greater than 1; then $|(a_{ij}^{-1}\cdot\zeta_j)^2|\leqq|\pi^2|$ since $|a_{ij}^{-1}|\leqq|\pi|$ and ζ_j^2 is integral. But $a_{ij}^{-1}\cdot\zeta_j$ is an element of L since its ξ_i coefficient is a unit; its norm is therefore a non-trivial representation of $\{\pi^2\}$ which is impossible. Hence a_{ij} must be integral for all i, j. Hence $K\subseteq L$ and since $|d(L)|=|d(K)|$, L must be equal to K.

We may now assume that both L and K represent $\{\pi^2\}$ non-trivially. First we complete the induction when n is even. Our assumptions guarantee a vector X in L such that $X^2=\{\pi^2\}$ and $X\cdot\xi_1$ (say) equals 1. By Lemma 4. 3 we may write $L=\mathfrak{o}\xi_1+\mathfrak{o}(\alpha\cdot X+\beta\cdot\xi_1)+\cdots$; then if $|\alpha|>1$ or if $|\beta|>1$, $|\alpha|=|\beta|$ since $\pi^{-1}\cdot X\notin L$ and $\pi^{-1}\cdot\xi_1\notin L$; now $\alpha\cdot X^2+\beta\cdot(X\cdot\xi_1)\,\varepsilon\,\mathfrak{o}$ and so $|\beta|\leqq 1$; hence $\alpha\,\varepsilon\,\mathfrak{o}$ also; it follows that $L=\mathfrak{o}\xi_1+\mathfrak{o}X+\cdots$ and hence $L=(\mathfrak{o}\xi_1+\mathfrak{o}X)\oplus L_1=L_0\oplus L_1$ in virtue of Lemma 2. 2. Then $L_0=\mathfrak{o}\Xi_1+\mathfrak{o}\Xi_2$ with $\Xi_2^2=1+2\beta$, $\beta\,\varepsilon\,\mathfrak{o}$, $\Xi_1^2=\{\pi^2\}$ and $\Xi_1\cdot\Xi_2=1$ in virtue of Lemma 4. 1. Then $\Xi_2-\beta\cdot\Xi_1\,\varepsilon\,L_0$ and its norm is equal to $1+\{\pi^2\}$. We now repeat this step and obtain $(\Xi_2-(\beta\pm\pi\gamma)\cdot\Xi_1)^2=1+\{\pi^3\}$. Hence by Lemma 4. 2 L_0 must represent 1. Therefore there is a $Y\,\varepsilon\,L$ and (by symmetry) a $Z\,\varepsilon\,K$. such that $Y^2=1=Z^2$. Then $L=\mathfrak{o}Y\oplus L_2$ and $K=\mathfrak{o}Z\oplus K_2$ in virtue of Lemmas 4. 3 and 2. 1. Now L_2 and K_2 are in isometric spaces by the Witt cancellation law; they are both proper being of unit determinant and of odd dimension; hence $L_2\cong K_2$; hence $L\cong K$.

Now let n be odd, $n\geqq 3$. Once again it is possible to write $L=L_0\oplus L_1$ with $L_1=\mathfrak{o}\Xi_3\oplus\cdots\oplus\mathfrak{o}\Xi_n$ and $L_0=\mathfrak{o}\Xi_1+\mathfrak{o}\Xi_2$ where $\Xi_1^2=\{\pi^2\}$, $\Xi_1\cdot\Xi_2=1$, $\Xi_2^2=1+\{\pi\}$. If $\Xi_1^2=[\pi^2]$, replace Ξ_1 by $\Xi_1+\gamma\pi\cdot\Xi_3$ in such a way that $(\Xi_1+\gamma\pi\cdot\Xi_3)^2=\{\pi^3\}$ (Lemma 4. 1) and write

$$L=(\mathfrak{o}(\Xi_1+\gamma\pi\cdot\Xi_3)+\mathfrak{o}\Xi_2)\oplus L_3$$

by Lemma 2. 2. In any case we have $L=L_2\oplus L_3$ and similarly $K=K_2\oplus K_3$ where L_2 and K_2 are proper lattices in 2-dimensional spaces of determinant $-1+\{\pi^3\}$, that is to say of determinant -1. These two spaces are isometric [20] and so $L_2\cong K_2$ since the theorem has already been shown to hold for $n=2$. It follows that L_1 and K_1 are in isometric spaces of odd dimension, are therefore both proper and hence by the inductive hypothesis are both isometric. Hence $L \cong K$. Q. E. D.

Note 4. 1. This is the first point at which the ramified theory differs essentially from the case where 2 is a prime, for in general it is not true that two totally integral lattices of unit determinant in the same metric space are isometric, even if their norms are the same. However, an answer can be given for this case and we will carry this out after finishing the unramified theory.

Note 4. 2. The restriction that the residue class field be perfect is also an important one and Lemma 4. 1 is fundamental in these discussions. Even if $e = 1$, Theorem 4. 4 need not hold with an imperfect residue class field as the following example will show: let Q be the rational numbers with the 2-adic valuation and extend the valuation to $Q[t]$ by defining $| a_0 + a_1 t + \cdot \cdot \cdot a_n t^n | = \max | a_\lambda |$ as in [1, p. 26] and then to $Q(t)$ by taking quotients. Let $Q(t)^*$ denote the completion of $Q(t)$. Then $(2t - 1) \cdot x^2 + (2t + 1) \cdot y^2$ and $t \cdot x^2 + (4t^2 - 1)/t \cdot y^2$ are fractionally equivalent but not integrally equivalent over $Q(t)^*$. If $e = 0$, the perfectness of the residue class field is no longer essential: see [3, 4] and Section 6 of this paper. The rational p-adic form of Theorem 4. 4 can be found in [9, Theorem 36].

5. The cancellation law. Examples [3] show that very little indeed can be expected in the way of a cancellation law over a general local field and this is especially true if 2 splits in F. However, the following form will be found useful in the ramified theory:

THEOREM 5. 1. *Let L be a lattice having the two decompositions $L = \mathfrak{o}\xi \oplus L_1$ and $L = \mathfrak{o}\eta \oplus K_1$ with $\xi^2 = \epsilon = \eta^2$ and $| \epsilon | = 1$. Suppose further that $| U \cdot W | \leqq | 2 |$ for all U and W in L_1. Then $L_1 \cong K_1$.*

Proof. Write $\eta = \epsilon_1 \cdot \xi + \pi^\gamma \cdot X$ with X a vector in L_1 having the properties $X \cdot \xi = 0$ and $\pi^{-1} \cdot X \notin L_1$. If $| X^2 | = \max | X \cdot Z |$, $Z \varepsilon L_1$, then

$$L = \mathfrak{o}\xi \oplus (\mathfrak{o}X + \textstyle\sum \mathfrak{o}T_\lambda) = \mathfrak{o}\xi \oplus \mathfrak{o}X \oplus \sum \mathfrak{o}T_\lambda' = \mathfrak{o}\eta \oplus \mathfrak{o}X' \oplus \sum \mathfrak{o}T_\lambda'$$

using lemmas 4. 3, 2. 1 and 2. 1 in that order. A determinantal consideration shows that $(X)^2 \cong (X')^2$. But $L_1 = \mathfrak{o}X \oplus \sum \mathfrak{o}T_\lambda'$ and $K_1 = \mathfrak{o}X' \oplus \sum \mathfrak{o}T_\lambda'$ and so $L_1 \cong K_1$.

So assume that $| X^2 | < \max | X \cdot Z |$, $Z \varepsilon L_1$, from now on. We will be interested in the number $\pi^{2\gamma} \cdot X^2 = \epsilon(1 - \epsilon_1)(1 + \epsilon_1)$. Either $(1 + \epsilon_1) \notin (2\pi) \cdot \mathfrak{o}$ or $(1 - \epsilon_1) \notin (2\pi) \cdot \mathfrak{o}$: for instance, $1 + \epsilon_1 \varepsilon (2\pi) \cdot \mathfrak{o}$ implies $1 - \epsilon_1 = 1 + \epsilon_1$

$-2\epsilon_1 = [2]$. By replacing ξ by $-\xi$ if necessary we can assume that $(1+\epsilon_1) \notin (2\pi)\cdot\mathfrak{o}$, that is $|1+\epsilon_1| \geqq |2|$.

Let $N \varepsilon L_1$ be such that $|X \cdot N| = \max |X \cdot Z|$, $Z \varepsilon L_1$. Then $|X^2| < |N \cdot X|$ and so X and N are linearly independent. Applying Lemma 4.3 to L_1 permits the expansion $L = \mathfrak{o}\xi \oplus (\mathfrak{o}X + \mathfrak{o}(\alpha \cdot X + \beta \cdot N) + \sum \mathfrak{o}T_\lambda)$. Suppose if possible that $|\beta| > 1$ or $|\alpha| > 1$ or both; then $\pi^{-1} \cdot X \notin L_1$ and $\pi^{-1} \cdot N \notin L_1$ and hence $|\alpha| = |\beta|$; but $|(\alpha \cdot X + \beta \cdot N) \cdot X| = |\beta| \cdot |N \cdot X| \leqq |N \cdot X|$ by choice of N; hence $|\beta| \leqq 1$ and so $|\alpha| \leqq 1$. Therefore we can write $L = \mathfrak{o}\xi \oplus (\mathfrak{o}X + \mathfrak{o}N + \sum \mathfrak{o}T_\lambda)$. Since $|X \cdot N| \geqq |X \cdot T_\lambda|$, it is possible to replace the T_λ by $(T_\lambda - (X \cdot T_\lambda)/(N \cdot X) \cdot N)$ and if we continue to call these new vectors T_λ, it follows that

$$(5.1) \qquad X \cdot T_\lambda = 0 \quad \text{and} \quad \xi \cdot T_\lambda = 0.$$

Then $L = \mathfrak{o}\xi + \mathfrak{o}X + \mathfrak{o}N + \sum \mathfrak{o}T_\lambda = \mathfrak{o}\eta + \mathfrak{o}(\epsilon_1^{-1} \cdot X) + \mathfrak{o}N + \sum \mathfrak{o}T_\lambda$ since $\eta = \epsilon_1\xi + \pi^\gamma \cdot X$ and since $\mathfrak{o}X = \mathfrak{o}(\epsilon_1^{-1} \cdot X)$. If we now use Lemma 2.1 to orthogonalize with respect to $\mathfrak{o}\eta$ we obtain

$$(5.2) \qquad L = \mathfrak{o}\eta \oplus (\mathfrak{o}Y + \mathfrak{o}M + \sum \mathfrak{o}T_\lambda)$$

where $Y = \epsilon_1^{-1}(X - (X \cdot \eta) \cdot \epsilon^{-1} \cdot \eta)$ and $M = (N - (N \cdot \eta) \cdot \epsilon^{-1} \cdot \eta)$ and direct computation yields the following equalities:

$$(5.3) \qquad \begin{gathered} Y \cdot T_\lambda = 0 = X \cdot T_\lambda;\ Y^2 = X^2;\ M \cdot Y = \epsilon_1 \cdot (N \cdot X);\ M \cdot T_\lambda = N \cdot T_\lambda; \\ M^2 = N^2 - \pi^{2\gamma}(N \cdot X)^2 \cdot \epsilon^{-1}. \end{gathered}$$

Now write $L = \mathfrak{o}\eta \oplus (\mathfrak{o}Y + \mathfrak{o}(M + \psi \cdot Y) + \sum \mathfrak{o}T_\lambda)$ where ψ is defined by

$$(5.4) \qquad \psi = \frac{\pi^{2\gamma}(N \cdot X)}{\epsilon(1+\epsilon_1)}.$$

Then ψ is an integer since $|N \cdot X| \leqq |2|$ and $|1+\epsilon_1| \geqq |2|$. Also $(M + \psi Y)^2 = N^2$, $(M + \psi Y) \cdot Y = N \cdot X$ and $(M + \psi Y) \cdot T_\lambda = N \cdot T_\lambda$ is easily shown by using equations (5.3) and (5.4). Also $X^2 = Y^2$, $X \cdot T_\lambda = 0 = Y \cdot T_\lambda$ and $T_\lambda \cdot T_\mu = T_\lambda \cdot T_\mu$. Hence L_1 and K_1 are isometric. Q. E. D.

Remark. Let us examine the previous proof under the following set of circumstances: $e \leqq 1$, L_1 and K_1 are totally integral and of the same type, so that $N(L_1) = N(K_1)$. If $|X^2| = \max |X \cdot Z|$, $Z \varepsilon L_1$, the proof goes through. If $|N \cdot X| \leqq |2|$, the proof goes through. So let us consider the remaining possibility which fails since ψ in equation (5.4) need not be integral: $|X^2| < \max |X \cdot Z|$ and $|N \cdot X| > |2|$. Then $|X^2| < 1$ and $|N \cdot X| = 1$ since L_1 is totally integral and F is unramified. The equations of Lemma 2.2 are therefore seen to have integral solutions when applied to $\mathfrak{o}X + \mathfrak{o}N + \sum \mathfrak{o}T_\lambda$. Hence $L = \mathfrak{o}\xi \oplus (\mathfrak{o}X + \mathfrak{o}N) \oplus L_2$. Now write $L_2 = L_3(\pi^0) \oplus L_4$ as in equation (2.1). We have therefore expressed $L = \mathfrak{o}\xi \oplus L_5 \oplus L_4$ where $X \varepsilon L_5$ and $|U \cdot W| \leqq |2|$ all U, W in L_4. Then $\mathfrak{o}\xi \oplus L_5 = \mathfrak{o}\eta + L_5 = \mathfrak{o}\eta \oplus K_5$ in virtue of Lemma 2.1. An application of the cancellation law for fractional equivalence shows that L_5 and K_5 are in isometric spaces. But they are also totally integral lattices of unit determinant and of the same type since $N(L_1) = N(K_1)$. Theorem 4.4 says they are isometric. We can therefore prove the following result which has already been established over the rational p-adic integers by Jones and Durfee [3, 7, 9, 10]:

THEOREM 5.2. *Let* $e \leqq 1$. *If* L *is a totally integral lattice having the two decompositions* $L = L_0 \oplus L_1 = K_0 \oplus K_1$ *in which* L_0 *and* K_0 *are isometric and of unit determinant, then* L_1 *and* K_1 *are isometric provided that* $N(L_1) = N(K_1)$.

Proof. The case where L_0 and K_0 are proper follows immediately by induction on $\dim L_0$ from the previous remark. So now assume that L_0 and K_0 are both improper. Construct a new metric space $V \oplus F \cdot \zeta$ with $\zeta^2 = 1$. Then $L^* = (\mathfrak{o}\zeta \oplus L_0) \oplus L_1 = (\mathfrak{o}\zeta \oplus K_0) \oplus K_1$ has proper first components and so $L_1 \cong K_1$. Q. E. D.

COROLLARY 5.2. *If* L_0 *and* K_0 *are isometric lattices, but not necessarily of unit determinant, and if the least divisor of* L_0 *is greater than or equal to the greatest divisor of* L_1, *and if* $N(L_1) = N(K_1)$, *then* $L_1 \cong K_1$.

6. The odd case. The theory of equivalence over fields for which $e = 0$ can be made to follow very rapidly at this stage. This has been completely worked out [3] so we merely state the results here. We use the notation of equation (2.1).

THEOREM 6.1. L *and* K *are isometric if and only if they are of the same type and* $L_i \cong K_i$ *for all* i, $1 \leqq i \leqq r$. *The latter condition is equivalent to* $\det(L_i) \cong \det(K_i)$ *over a local field whose residue class field is finite.*

THEOREM 6.2. *If* $L = L_0 \oplus L_1 = K_0 \oplus K_1$, *then* $L_1 \cong K_1$, *provided that* $L_0 \cong K_0$.

7. The space invariants. We are now closer to our goal of obtaining a complete set of invariants under integral equivalence and soon we will appeal for assistance to the known theory of fractional equivalence. In particular we will assume (using the language of metric spaces) that V and V' are isometric if and only if $n(V)=n(V')$, $d(V)=d(V')$ and $S(V)=S(V')$. The notation is as in [20]: n is the dimension of V, d the determinant of V (but for a square factor) and S the algebra $\prod_{\lambda\leqq\mu}(\xi_\lambda{}^2,\xi_\mu{}^2)$ where (ξ_λ) is any orthogonal basis for V. We will show that these invariants apply to local fields with perfect residue class fields of characteristic 2: the proof is not relevant to these discussions on integral equivalence so it is relegated to the appendix.

8. Two cases have to be discussed: (1) 2 is a prime in F; (2) 2 is ramified in F. The second requires a more delicate argument than the first, so we will treat the unramified theory separately; in any case we can say much more about it. But first we make the following remark in order to facilitate working with the S-symbols.

Remark. If F is a local field for which $e\geqq 1$, then $(1+4\alpha,\epsilon)=1$ if ϵ is a unit and $\alpha\in\mathfrak{o}$. Also $(1+4\alpha,\epsilon\pi)=1$ if and only if $1+4\alpha$ is a square in F.

Proof. The first part follows immediately if we note that $(1+4\alpha)x^2+\epsilon y^2$ always represents 1. The second follows if we observe that $(1+4\alpha)x^2+\epsilon\pi y^2$ does not represent 1 unless $(1+4\alpha)$ is a square.

II. The Unramified Theory.

Throughout this chapter we will assume that $|\pi|=|2|$.

9. Continuous lattices. Let us consider the canonical form $L=L_1(\pi^{s(1)})\oplus\cdots\oplus L_r(\pi^{s(r)})$ obtained in equation (2.1). We saw in Section 6 that over certain local fields the L_i are independent in the following sense: if $K_1(\pi^{s(1)})\oplus\cdots\oplus K_r(\pi^{s(r)})$ is another canonical form for L, then $L_i(\pi^{s(i)})\cong K_i(\pi^{s(i)})$. This is not the case when $|2|<1$ unless the differences $s(i+1)-s(i)$ are large enough; otherwise the L_i become interwoven and we shall then have to study larger blocks than the mere L_i. To this end we make the following definition:

Definition 9.1. L is said to be a continuous lattice if: (a) $s(i+1)=s(i)+1$, for all i satisfying $1\leqq i\leqq r-1$; (b) the L_i are proper for all $2\leqq i\leqq r$; (c) L_1 is either proper or improper. If L_1 is proper we say that L is entirely proper.

This notion is well-defined by Theorem 3.1.

THEOREM 9.1. *If L and K are entirely proper continuous lattices in V and if they are both of the same type, then $L \cong K$.*

Proof. We can take $\pi^{s(1)} = \pi^0$ and we can assume that $r > 1$ by Theorem 4.4. First let $n = 2$. Then there is an X in L such that $X^2 = 1 + \lambda\pi$, $\lambda \varepsilon \mathfrak{o}$, in virtue of Lemma 4.1. We want an X with $\lambda \varepsilon (\pi) \cdot \mathfrak{o}$. To that end assume that the λ found above is a unit and let Y be orthogonal to X and such that $Y^2 = \lambda\pi + \{\pi^2\}$; by Lemma 4.1 such a Y must exist. Then $(X + Y)^2 = 1 + \{\pi^2\}$. We can repeat the argument for K. Then $L = \mathfrak{o}Z_1 \oplus \mathfrak{o}Z_2$ and $K = \mathfrak{o}U_1 \oplus \mathfrak{o}U_2$ with $Z_1{}^2 = 1 + \alpha\pi^2$ and $U_1{}^2 = 1 + \beta\pi^2$, both α and β being in $\mathfrak{o}$. Apply the S-symbols to L:

$$S(L) \sim (1 + \alpha\pi^2, -d(L)) \cdot (d(L), d(L)).$$

Then $1 \sim S(L) \cdot S(K) \sim ((1 + \alpha\pi^2)(1 + \beta\pi^2), -d(L))$. Now $|-d(L)| = |\pi|$, hence $(1 + \alpha\pi^2)^{\frac{1}{2}} \cdot (1 + \beta\pi^2)^{\frac{1}{2}} \varepsilon F$ by the remark of Section 8. Hence $(1 + \alpha\pi^2) \cong (1 + \beta\pi^2)$ and therefore $L \cong K$.

For $n > 2$ it is convenient to separate the following two cases:

Case 1. $r = 2$, $\dim L_1 = 1$: Then $\dim L_2 > 1$. Write $L = \mathfrak{o}\xi_1 \oplus (\mathfrak{o}\xi_2 \oplus \cdots \oplus \mathfrak{o}\xi_n)$ with $\xi_2{}^2 = \pi(1 + \{\pi\})$ and then replace ξ_2 by $\Xi_2 = \xi_2 + \alpha\pi \cdot \xi_1$ in such a way that $\Xi_2{}^2 = \pi(1 + \{\pi^2\})$: that this is possible follows from Lemma 4.1. Now replace ξ_1 by $\Xi_1 = (\xi_1 - (\xi_1 \cdot \Xi_2) \cdot (\Xi_2{}^2)^{-1} \cdot \Xi_2)$; then $L = \mathfrak{o}\Xi_1 \oplus \mathfrak{o}\Xi_2 \oplus \mathfrak{o}\xi_3 \oplus \cdots \oplus \mathfrak{o}\xi_n$ since $(\xi_1 \cdot \Xi_2) \cdot (\Xi_2{}^2)^{-1} \varepsilon \mathfrak{o}$. Applying Lemma 4.1 again, we see that Ξ_2 can be replaced by $\Xi_2 + \beta\pi\xi_3$ so that $(\Xi_2 + \beta\pi\xi_3)^2 = \pi(1 + \{\pi^3\})$. Then in virtue of Lemmas 4.2 and 2.1 we see that $L = \mathfrak{o}X \oplus L_1$ with $X^2 = \pi$ and similarly $K = \mathfrak{o}Y \oplus K_1$ with $Y^2 = \pi$. Then L_1 and K_1 are still entirely proper and so $L_1 \cong K_1$ by induction and hence $L \cong K$.

Case 2. $r \geqq 3$ or $\dim L_1 > 1$: As in the case $n = 2$ we can write $L = \mathfrak{o}Z_1 \oplus \cdots \oplus \mathfrak{o}Z_n$ with $Z_1{}^2 = 1 + \{\pi^2\}$. Since $r \geqq 3$ or $\dim L_1 > 1$, there is a Z_i $(i \geqq 2)$ such that $(Z_1 + \alpha Z_i)^2 = 1 + \{\pi^3\}$ with $|\alpha| = 1$ or $|\alpha| = |\pi|$. We can therefore write $L = \mathfrak{o}(Z_1 + \alpha Z_i) \oplus L_1$ in virtue of Lemma 2.1, where L_1 is still entirely proper; and using Lemma 4.2, $L = \mathfrak{o}Z \oplus L_1$ with $Z^2 = 1$. And similarly $K = \mathfrak{o}W \oplus K_1$ with $W^2 = 1 = Z^2$ and K_1 entirely proper. Then L_1 and K_1 are in isometric spaces of dimension $n - 1$; by induction $L_1 \cong K_1$; hence $L \cong K$. Having considered all possible cases, the proof is now through.

THEOREM 9.2. *Any two continuous lattices of the same type and in the same metric space V are isometric.*

Proof. Again take $\pi^{s(1)} = \pi^0$ and use induction to n. We may take $r > 1$ by Theorem 4.4. Express L in a canonical basis: $L = (\mathfrak{o}\xi_1 + \mathfrak{o}\xi_2) \oplus L_1$. If $F \cdot \xi_1 + F \cdot \xi_2$ represents 0 non-trivially, do not proceed further. If not, then $\xi_1{}^2 = \alpha\pi$, α being a unit. Since $r > 1$, we can choose an X orthogonal to ξ_1 in such a way that $X^2 = \alpha\pi + \{\pi^2\}$ in virtue of Lemma 4.1. Then write $L = (\mathfrak{o}(\xi_1 + X) + \mathfrak{o}\xi_2) \oplus L_2 = L_0 \oplus L_2$, which can be done by Lemma 2.2. Moreover L_0 represents zero non-trivially since $d(L_0) \cong -1$. Similarly for $K = K_0 \oplus K_2$. Then $L_0 \cong K_0$ by Theorem 4.4. So L_2 and K_2 are in isometric spaces and hence $L_2 \cong K_2$. Therefore $L \cong K$.

10. Congruence of lattices. Let us prove some fundamental lemmas which we will need in the next section. By an i-dimensional lattice in V we mean a subset of V that is a lattice in some i-dimensional subspace of V.

Definition 10.1. If L and K are two i-dimensional lattices of V, we say that $L \cong K \pmod{\pi^\omega}$ if there is a transformation A such that $AV = V$, $AL = K$ and $AX \cdot AY \equiv X \cdot Y \pmod{\pi^\omega}$ for all X and Y in L.

We will not need this definition until much later. We mention it here because of the analogy between it and the sharper conditions (2) of the next two lemmas.

Now let L have the matrix $\begin{pmatrix} \{\pi\} & 1 \\ 1 & \{\pi\} \end{pmatrix}$. Then by simple changes of basis and using Lemma 4.1 we have $L \cong \begin{pmatrix} [\pi] & 1 \\ 1 & \{\pi\} \end{pmatrix} \cong \begin{pmatrix} \pi + \{\pi^2\} & 1 \\ 1 & \alpha\pi \end{pmatrix}$ with $\alpha \in \mathfrak{o}$. Put $J \cong \begin{pmatrix} \pi & 1 \\ 1 & \alpha\pi \end{pmatrix}$. Contention: $L \cong J$. For if $\alpha \in (\pi) \cdot \mathfrak{o}$ then L and J are in spaces of determinant (-1), hence in isometric spaces and hence $L \cong J$ by Theorem 4.4. So now assume that α is a unit. Then $d(L) \cong d(K)$. Also $S(L) \sim (\alpha\pi, \alpha\pi) \cdot (\alpha\pi \cdot d(L), d(L)) \sim S(J)$. Hence L and K are in isometric spaces and so once again $L \cong J$.

LEMMA 10.1. *If L and K are two i-dimensional lattices of V, both having the same least divisor $|\pi^m|$, and if there is a transformation A of V onto itself with the properties*: (1) $AL = K$; (2) *if X and $Y \in L$, then* $AX \cdot AY \equiv X \cdot Y \pmod{\pi^{m+1}}$ *and* $(AX)^2 \equiv X^2 \pmod{\pi^{m+2}}$; (3) $d(L) \cong d(K)$; *then L is isometric to K.*

Proof. By induction to i. We use the notation of equation (2.1). There is no loss in generality if we assume that $s(1) = 0$ for L and that both L and K are totally integral. If $i = 1$ there is nothing to prove. If $i = 2$ and L is improper, then $L \cong \begin{pmatrix} \pi & 1 \\ 1 & \alpha\pi \end{pmatrix}$ by the contention proved above and so $K \cong \begin{pmatrix} \pi + \{\pi^2\} & 1 + \{\pi\} \\ 1 + \{\pi\} & \alpha\pi + \{\pi^2\} \end{pmatrix} \cong \begin{pmatrix} \pi + \{\pi^2\} & 1 \\ 1 & \alpha\pi + \{\pi^2\} \end{pmatrix}$. If $\alpha \varepsilon (\pi) \cdot \mathfrak{o}$, then $d(L) \cong -1 \cong d(K)$ and Theorem 4. 4 carries us through. If $|\alpha| = 1$, the result follows from the contention proved above.

We can now start the induction for general i, but distinguishing the three cases:

Case 1. $r = 1$ and $L_1(\pi^{s(1)})$ is proper: Write $L = \mathfrak{o}\xi_1 \oplus \cdots \oplus \mathfrak{o}\xi_i = \mathfrak{o}\xi_1 \oplus L_2$ and then $K = \mathfrak{o}(A\xi_1) + \cdots + \mathfrak{o}(A\xi_i) = \mathfrak{o}\eta_1 \oplus (\mathfrak{o}\eta_2 + \cdots + \mathfrak{o}\eta_i) = \mathfrak{o}\eta_1 \oplus K_2$, where $\eta_1 = A\xi_1$ and $\eta_j = A\xi_j - (A\xi_1 \cdot A\xi_j) \cdot ((A\xi_1)^2)^{-1} \cdot A\xi_1$ for $2 \leqq j \leqq i$. Then it is easy to see that $\eta_j^2 \equiv \xi_j^2 \pmod{\pi^2}$ and that $\xi_j \cdot \xi_h \equiv \eta_j \cdot \eta_h \pmod{\pi}$. Now write $K = \mathfrak{o}(\eta_1 + \alpha\pi \cdot \eta_2) + \mathfrak{o}\eta_2 + \cdots + \mathfrak{o}\eta_i$ with α chosen so that $(\eta_1 + \alpha\pi \cdot \eta_2)^2 \cong \xi_1^2$: this can be done in virtue of Lemmas 4. 1 and 4. 2, observing that $\eta_1^2 = (A\xi_1)^2 \equiv (\xi_1^2) \pmod{\pi^2}$. Now use Lemma 2. 1 to write $K = \mathfrak{o}(\eta_1 + \alpha\pi \cdot \eta_2) \oplus K_2^*$. Then K_2^* and L_2 have the properties enunciated in the statement of the lemma, hence $L_2 \cong K_2^*$ and so $L \cong K$.

Case 2. $r \geqq 1$, $L_1(\pi^{s(1)})$ improper: Express L canonically:

$$L = (\mathfrak{o}\xi_1 + \mathfrak{o}\xi_2) \oplus (\mathfrak{o}\xi_3 + \cdots + \mathfrak{o}\xi_i) = L_0 \oplus L_1.$$

Then it is possible to write

$$K = (\mathfrak{o}(A\xi_1) + \mathfrak{o}(A\xi_2)) \oplus (\textstyle\sum \mathfrak{o}(A\xi_\lambda + \alpha_\lambda(A\xi_1) + \beta_\lambda(A\xi_2))) = K_0 \oplus K_1$$

where α_λ and β_λ are determined by Lemma 2. 2 and are therefore in $(\pi^{m+1}) \cdot \mathfrak{o}$. Then $L_0 \cong K_0$ and we can apply induction to L_1 and K_1 to obtain $L \cong K$.

Case 3. $r > 1$, $L_1(\pi^{s(1)})$ proper: This is almost the same as Case 1, except that $m > 0$. We get $K = \mathfrak{o}\eta_1 \oplus K_2$ as before. Then $\eta_1^2 \equiv \xi_1^2 \pmod{\pi^3}$, hence $\eta_1^2 \cong \xi_1^2$, hence $L \cong K$. Q. E. D.

Definition 10. 2. If $L_1(\pi^{s(1)}) \oplus \cdots \oplus L_r(\pi^{s(1)})$ is a canonical form for L, we say that L is a proper lattice if $L_r(\pi^{s(r)})$ is proper; we say that L is improper if $L_r(\pi^{s(r)})$ is improper.

The reason for this strange definition will become apparent in the next two sections.

LEMMA 10. 2. *If L and K are two i-dimensional lattices in V, both having the same least divisor $|\pi^m|$, and if there is a transformation A of V onto itself with the properties*: (1) $AL = K$; (2) $AX \cdot AY \equiv X \cdot Y \pmod{\pi^\omega}$ *and* $(AX)^2 \equiv X^2 \pmod{\pi^{\omega+1}}$ *for all X and Y in L, where $\omega = m + 1$ if L is improper and $\omega = m + 2$ if L is proper, then L and K are isometric.*

Proof. An inductive argument yields $d(L) \cong d(K)$. Now apply Lemma 10. 1. Q. E. D.

11. Discontinuous decompositions. In Section 9 we pointed out that the $L_i(\pi^{s(i)})$ would be too small for our purposes and we accordingly introduced larger blocks which we called continuous lattices. It is clear how to subdivide a lattice $L = L_1(\pi^{s(1)}) \oplus \cdots \oplus L_r(\pi^{s(r)})$ into continuous lattices of maximal dimension. Let us write the subdivision $L = M(1) \oplus \cdots \oplus M(q)$. If $L = M'(1) \oplus \cdots \oplus M'(q)$ is another such decomposition, it is clear that $M(i)$ will have the same type as $M'(i)$, but it is still not true that $M(i)$ must be isometric to $M'(i)$. However, it will not be difficult to show how the $M(i)$ change when we pass to the $M'(i)$. To that end we make the following definition:

Definition 11.1. Consider the decomposition $L = K \oplus J$. We say that this decomposition is discontinuous if either (1) K is improper and $(\pi^{m(K)+2}) \cdot \mathfrak{o} \supseteq N(J)$ or (2) K is proper and $(\pi^{m(K)+3}) \cdot \mathfrak{o} \supseteq N(J)$. We call the decomposition semicontinuous if (3) K is proper and $(\pi^{m(K+2)}) \cdot \mathfrak{o} = N(J)$. The notation is an in Definitions 3.1 and 10.2.

LEMMA 11.1. *Let L and K be two lattices with discontinuous decompositions $L = L(1) \oplus L(2)$ and $K = K(1) \oplus K(2)$ such that $L(1)$ and $K(1)$ have the same type. Then $L \cong K$ if and only if $L(1) \cong K(1)$ and $L(2) \cong K(2)$.*

Proof. The sufficiency is clearly true. To prove the necessity we can assume that $L = K$ and then

$$L = L(1) \oplus L(2) = (\mathfrak{o}\xi_1 + \cdots + \mathfrak{o}\xi_i) \oplus (\mathfrak{o}\xi_{i+1} + \cdots + \mathfrak{o}\xi_n)$$

$$L = K(1) \oplus K(2) = (\mathfrak{o}\eta_1 + \cdots + \mathfrak{o}\eta_i) \oplus (\mathfrak{o}\eta_{i+1} + \cdots + \mathfrak{o}\eta_n)$$

in the canonical bases (ξ) and (η). Then

$$\begin{array}{l} \eta_1 = a_{11}\xi_1 + \cdots + a_{i1}\xi_i + N_1 = \Xi_1 + N_1 \\ \cdots \qquad \cdots \qquad \cdots \qquad \cdots \\ \eta_i = a_{1i}\xi_1 + \cdots + a_{ii}\xi_i + N_i = \Xi_i + N_i \end{array} \tag{11.1}$$

where $a_{ij} \in \mathfrak{o}$ since $\eta_j \in L$. Also $|N_j^2| \leqq |\pi^{\omega+1}|$ by Definition 11.1 and similarly $|N_j \cdot Z| \leqq |\pi^{\omega}|$ all $Z \in L$, ω being $m(L(1)) + 1$ or $m(L(1)) + 2$, depending on whether $L(1)$ is improper or proper.

Now it is clear from equation (11.1) that $|\det(\mathfrak{o}\Xi_1 + \cdots + \mathfrak{o}\Xi_i)| = |\det(K(1))| = |\det(L(1))|$. Hence $\mathfrak{o}\Xi_1 + \cdots + \mathfrak{o}\Xi_i = L(1)$. Define $A\Xi_j = \eta_j$ for all $j \leqq i$ and extend it to a transformation of V onto V. Then $\eta_j^2 \equiv \Xi_j^2 \pmod{\pi^{\omega+1}}$ and $\eta_j \cdot \eta_h \equiv \Xi_j \cdot \Xi_h \pmod{\pi^{\omega}}$; hence $(AX)^2 \equiv X^2 \pmod{\pi^{\omega+1}}$ and $AX \cdot AY \equiv X \cdot Y \pmod{\pi^{\omega}}$ for all X and Y in $L(1)$. The conditions of Lemma 10.2 are satisfied and so $L(1) \cong K(1)$. Applying Corollary 5.2 yields $L(2) \cong K(2)$ and the lemma is proved.

LEMMA 11.2. *Let the lattice L have two semi-continuous decompositions $L = L(1) \oplus L(2)$ and $L = K(1) \oplus K(2)$. Then $d(L(1))/d(K(1)) \cong 1 \pmod{\pi^2}$ and*

$$(d(L(1)), \pi^{\operatorname{ord} d(L(1)) + m(L(1))}) \cdot S(L(1))$$
$$\sim (d(K(1)), \pi^{\operatorname{ord} d(K(1)) + m(K(1))}) \cdot S(K(1))$$

provided that $L(1)$ and $K(1)$ are both of the same type.

Proof. We obtain the system of equations (11.1) in the same way as before. Then $a_{ij} \varepsilon \mathfrak{o}$ and $|N_j^2| \leqq |\pi^{m+2}|$ and similarly $|N_j \cdot Z| \leqq |\pi^{m+1}|$ all Z in L. Once again $L(1) = \mathfrak{o}\Xi_1 + \cdots + \mathfrak{o}\Xi_i$. In addition it is easy to see that $d(L(1))/d(K(1)) \cong 1 \pmod{\pi^2}$ by considering the determinant of $(\Xi_\lambda \cdot \Xi_\mu)$ where $\Xi_\lambda = \eta_\lambda - N_\lambda$.

Now suppose for the moment that $d(L(1)) \cong d(K(1))$. Define $A\Xi_j = \eta_j$ for $j \leqq i$ and extend it to V. The conditions of Lemma 10.1 are satisfied and so $L(1) \cong K(1)$, hence

$$S(L(1)) \sim S(K(1)) \tag{11.2}$$

and hence the lemma is true.

If on the other hand $d(L(1))/d(K(1)) = a^2 + \epsilon\pi^2$, ϵ being a unit, it is possible to write

$$L = (\mathfrak{o}\xi_1 + \cdots + \mathfrak{o}\xi_{i-1} + \mathfrak{o}(\xi_i + \alpha\xi_{i+1})) \oplus L(2)^* = L(1)^* \oplus L(2)^*, \tag{11.3}$$

$\alpha \varepsilon \mathfrak{o}$, where ξ_{i+1} has norm $|\xi_{i+1}^2| = |\pi^{m+2}|$ in virtue of Lemma 2.1. Now ξ_i and therefore $(\xi_i + \alpha\xi_{i+1})$ is orthogonal to all the remaining $\xi_1, \cdots, \xi_{i-1}$ since the original decomposition was semi-continuous. Hence, with the aid of Lemma 4.1, an $\alpha \varepsilon \mathfrak{o}$ can be chosen in such a way that $d(L(1)^*)/d(K(1)) = a^2 + \{\pi^3\}$ and therefore $d(L(1)^*) \cong d(K(1))$.

Referring to equation (11.2) we see that

$$S(L(1)) \cdot S(K(1)) \sim S(L(1)) \cdot S(L(1)^*).$$

Direct computation shows that

$$S(L(1)) \cdot S(L(1)^*) \sim (-d(\mathfrak{o}\xi_1 + \cdots + \mathfrak{o}\xi_{i-1}), d(L(1)^*) \cdot d(L(1)))$$

and this is the same as

$$(\pi^{\operatorname{ord} d(L(1)) + m(L(1))}, d(L(1)) \cdot d(K(1)))$$

since

$$d(L(1))d(L(1)^*) \cong d(L(1))d(K(1)) \cong 1 + \{\pi^2\},$$

in virtue of the remark of Section 8. Hence the second part of the lemma is generally true.

12. Complete set of invariants. We now have enough available to be able to tell whether two lattices are isometric or not and we tie these results together in the form of a complete set of invariants. We will work with the decomposition $L = M(1) \oplus \cdots \oplus M(q)$ given in Section 11. Let K be another lattice in V having the same type as L and with the decomposition $K = P(1) \oplus \cdots \oplus P(q)$.

If we consider the decomposition $L = L(1) \oplus L(2)$ where

$$L(1) = M(1) \oplus \cdots \oplus M(j) \text{ and } L(2) = M(j+1) \oplus \cdots \oplus M(q),$$

then it is easy to see that this decomposition is either discontinuous or else semi-continuous. Hence, if $L \cong K$ it follows from Lemmas 11.1 and 11.2 that

$$\text{(12.1)} \qquad \frac{d(M(1) + \cdots + M(j))}{d(P(1) + \cdots + P(j))} \cong 1 \pmod{\pi^{\gamma(j)}}$$

holds for any j, $1 \leqq j \leqq q$, $\gamma(j)$ being 2 or 3 depending on whether the break from j to $j+1$ is semi-continuous or discontinuous. In addition

$$\text{(12.2)} \quad (d(M(1) + \cdots + M(j)), \pi^{\operatorname{ord} d(M(1)+\cdots+M(j))+m(j)}) \cdot S(M(1) + \cdots + M(j))$$

is the same for both lattices, if $m(j)$ denotes the ordinal of the least divisor of $M(j)$.

We now show that the converse is true by using induction to q. Let

$$M = M(1) \oplus \cdots \oplus M(q-1) = \mathfrak{o}\xi_1 + \cdots + \mathfrak{o}\xi_i$$

and let

$$P = P(1) \oplus \cdots \oplus P(q-1) = \mathfrak{o}\eta_1 + \cdots + \mathfrak{o}\eta_i,$$

the bases being canonical. If the break from $q-1$ to q is discontinuous, then $d(P) \cong d(M)$ and $S(P) \sim S(M)$ and hence, by the induction, $P \cong M$. Then $P(q)$ and $M(q)$ are continuous lattices in isometric spaces and hence isometric by Theorem 9.2. Hence $L \cong K$.

If, however, the break from $q-1$ to q is semi-continuous, we must proceed with a little more caution. Exactly as in equation (11.3) we can change ξ_i into $\xi_i + \alpha\xi_{i+1}$ so that L now appears as

$$L = M(1) \oplus \cdots \oplus M(q-2) \oplus M(q-1)^* \oplus M(q)^* = M^* \oplus M(q)^*$$

in such a way that $d(M^*) \cong d(P)$. Then direct computation yields

$S(M) \cdot S(M^*) \sim S(M) \cdot S(P)$ and hence $M^* \cong P$, by the induction. Then $M(q)^* \cong P(q)$, being continuous lattices in isometric spaces. Hence $L \cong K$.

We can therefore formulate the invariants in

THEOREM 12. 1. *A complete set of invariants for L is given by*

(1) *the type of L*

(2) $\left(\dfrac{d(M(1) + \cdots + M(j))}{\pi^{\operatorname{ord} d(M(1)+\cdots+M(j))+\gamma(j)}}\right)$

(3) $(\pi^{\operatorname{ord} d(M(1)+\cdots+M(j))+m(j)}, d(M(1) + \cdots + M(j))) \cdot S(M(1) + \cdots + M(j))$

for all j, $1 \leqq j \leqq q$, $\gamma(j)$ being 2 at a semi-continuity and 3 at a discontinuity.

III. The Ramified Theory.

13. We now discuss the ramified case and we aim to show that two totally integral lattices L and K of unit determinant and in the same metric space V are isometric, provided that they represent the same numbers. The main difficulty will be in the case of improper lattices and our plan is to settle (1) the 2-dimensional case (2) the 4-dimensional improper case (3) the general improper case (4) the general case.

We observe that the problem is already solved in Theorem 4. 4 whenever $N(L) = (2) \cdot \mathfrak{o}$ and we accordingly introduce this notation: ν will denote the integer defined by $N(L) = (\pi^\nu) \cdot \mathfrak{o}$ and we will assume that $\nu < e$; unless otherwise stated we will make the further assumption that $0 < \nu < e$.

14. Two-dimensional lattices. Throughout this section L and K will stand for improper 2-dimensional lattices of unit determinant with norm π^ν where $0 < \nu < e$. We shall show how to reduce such lattices to a certain standard form.

LEMMA 14. 1. *Let $L = \mathfrak{o}\xi_1 + \mathfrak{o}\xi_2$ be an improper lattice of unit determinant such that $|\xi_1{}^2| = |\pi^\nu|$ and $\xi_1 \cdot \xi_2 = 1$. Then there is an X in L such that $L = \mathfrak{o}\xi_1 + \mathfrak{o}X$, $\xi_1 \cdot X = 1$ and $X^2 = [\pi^\mu]$ where μ $(> \nu)$ is such that either $\mu \geqq 2e - \nu$ or else $\mu + \nu$ is odd.*

Proof. Express $L = \mathfrak{o}\xi_1 + \mathfrak{o}Y$ with $\xi_1 \cdot Y = 1$ and such that $|Y^2|$ is least for all such Y. If $\operatorname{ord} Y^2 \geqq 2e - \nu$ or $\operatorname{ord} Y^2 + \nu$ is odd, we are through. So assume that $\operatorname{ord} Y^2 < 2e - \nu$ and that $\operatorname{ord} Y^2 + \nu$ is even. In virtue of Lemma 4. 1 there is a unit α such that

$$(\alpha\pi^r\xi_1)^2 \equiv Y^2 \pmod{\pi Y^2} \tag{14. 1}$$

where $r = \frac{1}{2}(\text{ord } Y^2 - \nu)$ and $\alpha\pi^r$ is of course integral. Put $X = Y + \alpha\pi^r\xi_1$ and then $L = \mathfrak{o}\xi_1 + \mathfrak{o}X$. Then $| X^2 | = | Y^2 + \alpha^2\pi^{2r}\xi_1{}^2 + 2\alpha\pi^r\xi_1 \cdot Y | < | Y^2 |$ by equation (14.1) and since $| 2\alpha\pi^r\xi_1 \cdot Y | < | Y^2 |$. Also $\xi_1 \cdot X = \delta$ is a unit. Hence $L = \mathfrak{o}\xi_1 + \mathfrak{o}(\delta^{-1}X)$ has the properties $\xi_1 \cdot (\delta^{-1}X) = 1$ and $| (\delta^{-1}X)^2 | < | Y^2 |$. This denies the minimality of $| Y^2 |$ and therefore proves the theorem. Q. E. D.

Remark 14.1. If the μ of Lemma 14.1 which corresponds to the lattice L is such that $\mu < 2e - \nu$ and if K is another lattice with $d(K) \equiv d(L) \pmod 4$ and $N(L) = N(K)$, then the $\mu(K)$ corresponding to K is equal to μ.

Proof. Let us assume that $\mu(K) > \mu$ if possible. Then $d(L)/d(K) = 1 + [\pi^{\mu+\nu}]$ is a square (mod 4) and so $1 + [\pi^{\mu+\nu}] \equiv (1 + \epsilon_1\pi + \epsilon_2\pi^2 + \cdots)^2 \pmod 4$; direct expansion shows that this is impossible if $\nu + \mu$ is odd and less than $2e$. Q. E. D.

Notation. Let c's denote fixed representatives in $\mathfrak{o}$ of the residue class field $\mathfrak{o}/(\pi)\cdot\mathfrak{o}$ with the convention that 0 shall be the representative of the zero residue class: thus $c' \equiv c'' \pmod \pi$ if and only if $c' = c''$.

LEMMA 14.2. *If $L = \mathfrak{o}\xi_1 + \mathfrak{o}\xi_2$ with $| \xi_1{}^2 | = | \pi^\nu |$, then there is an X in L such that $L = \mathfrak{o}\xi_1 + \mathfrak{o}X$, $\xi_1 \cdot X = 1$ and*

$$X^2 = (c_0 + c_2\pi^2 + \cdots + c_{2k}\pi^{2k})\pi^\mu + \{\pi^{2e-\nu}\}$$

where $2k + \mu + 1 = 2e - \nu$.

Proof. There is nothing to prove if $\mu \geqq 2e - \nu$; so assume that $\mu < 2e - \nu$. Suppose that we have achieved $c_1 = c_3 = c_5 = \cdots = c_{2i-1} = 0$ when L is written in the form

$$\mathfrak{o}\xi_1 + \mathfrak{o}Y\colon\ Y^2 = (c_0 + c_2\pi^2 + \cdots + c_{2i}\pi^{2i} + d\pi^{2i+1})\pi^\mu;\ \xi_1 \cdot Y = 1.$$

We aim to find a U such that $L = \mathfrak{o}\xi_1 + \mathfrak{o}U$, $U \cdot Y = 1$ and

$$U^2 = (c_0 + c_2\pi^2 + \cdots + c_{2i}\pi^{2i} + c_{2i+2}\pi^{2i+2} + \{\pi^{2i+3}\})\pi^\mu \tag{14.2}$$

provided that $2i + \mu + 1 < 2e - \nu$.

If $d \in (\pi)\cdot\mathfrak{o}$ we are through. If not, there is a unit α such that $(\alpha\pi^r\xi_1)^2 \equiv d\pi^{2i+\mu+1} \pmod{\pi^{2i+\mu+2}}$ where $r = \frac{1}{2}(2i + \mu + 1 - \nu)$ in virtue of Lemma 4.1. Put $L = \mathfrak{o}\xi_1 + \mathfrak{o}Z$ where $Z = Y + \alpha\pi^r\xi_1$. We contend that

$$Z^2 = (c_0 + c_2\pi^2 + \cdots + c_{2i}\pi^{2i} + \{\pi^{2i+2}\})\pi^\mu. \tag{14.3}$$

We have $Y^2 + \alpha^2\pi^{2r}\xi_1{}^2$ equal to the right hand side of equation (14.3) by choice of $\alpha\pi^r$; in addition $2i + \mu + 1 < 2e - \nu$ implies that

$$e + r > 2i + \mu + 1 \tag{14.4}$$

by definition of r; and so $2\alpha\pi^r\xi_1 \cdot Y \varepsilon (\pi^{2i+\mu+2}) \cdot \mathfrak{o}$. Hence equation (14. 3) holds.

Now write $U = Z/(\xi_1 \cdot Z)$ and $L = \mathfrak{o}\xi_1 + \mathfrak{o}U$. Then $\xi_1 \cdot U = 1$. In addition, equation (14. 2) is satisfied: for

$$U^2 = Z^2(1 + \{\pi^{\nu+r}\})^2 = Z^2(1 + \{\pi^{2\nu+2r}\} + \{\pi^{\nu+r+e}\})$$

and $2\nu + 2r \geqq 2i + \mu + 2$ and $\nu + r + e \geqq 2i + \mu + 2$, by equation (14. 4).

We have therefore shown how to pass from Y to U. If we repeat this process sufficiently often we arrive at the X mentioned in the statement of the lemma. Q. E. D.

THEOREM 14. 3. *Let L and K be two 2-dimensional lattices such that $N(L) = (\pi^\nu) \cdot \mathfrak{o} = N(K)$. If there is an X in L and a Y in K such that $X^2 = \epsilon\pi^\nu = Y^2$ and if in addition $d(L) \cong d(K)$, then L is isometric to K. The result also holds when $e = \nu$.*

Proof. Applying the invariants of Section 7 we see that L and K can be taken in the same metric space V. If $\nu = e$ an easy computation shows that L and K are maximal [11, p. 50], and the theorem follows from [11, Satz 9. 6]. So we can assume that $\nu < e$. In virtue of Lemmas 4. 3, 14. 1 and 14. 2 we can write $L = \mathfrak{o}\xi_1 + \mathfrak{o}\xi_2$ where $\xi_1{}^2 = \epsilon\pi^\nu$, $\xi_1 \cdot \xi_2 = 1$ and $\xi_2{}^2 = (c_0 + c_2\pi^2 + \cdots + c_{2k}\pi^{2k})\pi^\mu + \alpha(\pi^{2e-\nu})$. In virtue of Lemmas 4. 3, 14. 1, 14. 2 and the Remark 14. 1 we can write $K = \mathfrak{o}\eta_1 + \mathfrak{o}\eta_2$ where $\eta_1{}^2 = \epsilon\pi^\nu$, $\eta_1 \cdot \eta_2 = 1$ and $\eta_2{}^2 = (d_0 + d_2\pi^2 + \cdots + d_{2k}\pi^{2k})\pi^\mu + \beta(\pi^{2e-\nu})$. Contention: $d_{2i} = c_{2i}$ all i. If not, let j be the smallest integer for which $d_{2j} \neq c_{2j}$: then $| d_{2j} - c_{2j} | = 1$. Since the determinants differ by a square factor we have

$$\begin{aligned}(-1 + \delta\pi^{\nu+\mu} &+ \epsilon d_{2j}\pi^{\nu+\mu+2j}) \\ &\equiv (-1 + \delta\pi^{\nu+\mu} + \epsilon c_{2j}\pi^{\nu+\mu+2j}) \cdot (1 + A\pi)^2 \qquad (\operatorname{mod} \pi^{\nu+\mu+2j+1}).\end{aligned}$$

Hence

$$(14.5) \qquad | \pi^{\nu+\mu+2j} | = | (d_{2j} - c_{2j})\pi^{\nu+\mu+2j} | = | A^2\pi^2 + 2A\pi |.$$

The ordinal of the left hand side of equation (14. 5) is odd and less than $2e$; that is, $| A^2\pi^2 + 2A\pi | > | \pi^{2e}$ and therefore $| A^2\pi^2 + 2A\pi | = | A^2\pi^2 |$. Hence the right hand side of equation (14. 5) has an even ordinal. This contradiction establishes the contention.

Now we must deal with $\alpha\pi^{2e-\nu}$ and $\beta\pi^{2e-\nu}$. The α and β are integers; if $\alpha = \beta$, then $L \cong K$ and we are through. So assume that $\alpha \neq \beta$. Define an f such that $(\beta - \alpha)\pi^{-f}$ has ordinal either $-e$ or $-e - 1$, the choice being made in such a way that $f + \nu$ is even. Clearly $f \geqq e > \nu$. Define a new metric space $V^* = F\cdot\zeta \oplus V$ with $\zeta^2 = (\beta - \alpha)\pi^{-f}$ and put $L^* = \mathfrak{o}\zeta \oplus L$ and $K^* = \mathfrak{o}\zeta \oplus K$. According to Theorem 5. 1 it suffices to prove that $L^* \cong K^*$. Write $L^* = \mathfrak{o}\zeta + \mathfrak{o}\xi_1 + \mathfrak{o}\Xi_2$ with $\Xi_2 = \xi_2 + \pi^r\zeta$ and $r = \frac{1}{2}(2e + f - \nu)$. Then $r \geqq e + 1$. Also $(\mathfrak{o}\xi_1 + \mathfrak{o}\Xi_2) \cong (\mathfrak{o}\eta_1 + \mathfrak{o}\eta_2)$ by the contention proved above and by choice of ζ and r. If we apply Lemma 2. 2 we can write $L^* = \mathfrak{o}(\zeta + a\xi_1 + b\Xi_2) \oplus (\mathfrak{o}\xi_1 + \mathfrak{o}\Xi_2)$ since it can easily be checked that a and b are integers. A determinantal argument shows that $\zeta^2 \cong (\zeta + a\xi_1 + b\Xi_2)^2$. Hence $L^* \cong K^*$. Hence $L \cong K$. Q. E. D.

COROLLARY 14. 3. *If* $L = \mathfrak{o}\xi_1 + \mathfrak{o}\xi_2$ *with* $\xi_1{}^2 = a_1\pi^\nu$, $\xi_1\cdot\xi_2 = 1$, $\xi_2{}^2 = \{\pi^\gamma\}$, γ *being* $\geqq e + 1$, *and if* $K = \mathfrak{o}\eta_1 + \mathfrak{o}\eta_2$ *with* $\eta_1{}^2 = a_1\pi^\nu + \{\pi^e\}$, $\eta_1\cdot\eta_2 = 1$, $\eta_2{}^2 = \{\pi^\gamma\}$, *then* $L \cong K$ *provided that* $d(L) \cong d(K)$.

Proof. Since $\gamma \geqq e + 1$ there is an integer δ such that $(\eta_1 + \delta\eta_2)^2 = a_1\pi^\nu$. The result then follows from Theorem 14. 3.

LEMMA 14. 4. *If* $L \cong K$ (mod $\pi^{2e-\nu+1}$), *then* $L \cong K$. *This result also holds when* $\nu = e$.

Proof. Case 1. $\nu \neq e$: We can write $L = \mathfrak{o}\xi_1 + \mathfrak{o}\xi_2$ with $\xi_1{}^2 = \epsilon_1\pi^\nu$, $\xi_1\cdot\xi_2 = 1$, $\xi_2{}^2 = \epsilon_2\pi^\nu$ where ϵ_1 and ϵ_2 are units. Since $K \cong L\,(\text{mod}\,\pi^{2e-\nu+1})$, it is easy to see that K can be written as $\mathfrak{o}\eta_1 + \mathfrak{o}\eta_2$ with $\eta_1{}^2 = \epsilon_1\pi^\nu + \{\pi^{2e-\nu+1}\}$, $\eta_1\cdot\eta_2 = 1$, $\eta_2{}^2 = \epsilon_2\pi^\nu + \{\pi^{2e-\nu+1}\}$. Successive application of Theorem 14. 3 and Hensel's lemma yields

$$K \cong \begin{pmatrix} \epsilon_1\pi^\nu + \{\pi^{2e-\nu+1}\} & 1 \\ 1 & \epsilon_2\pi^\nu + \{\pi^{2e-\nu+1}\} \end{pmatrix} \cong \begin{pmatrix} \epsilon_1\pi^\nu + \{\pi^{2e-\nu+1}\} & 1 \\ 1 & \epsilon_2\pi^\nu \end{pmatrix} \cong L.$$

Case 2. $\nu = e$: The same argument applies if $d(L) \not\cong -1$. If $d(L) \cong -1$, then $d(K) \cong -1$, and the result follows from the result on maximal lattices [11].

15. Congruence of lattices.

Although we will not have occasion to refer to it in this paper (in view of the sharper results at our disposal for lattices of unit determinant) we mention here a useful application of Section 14 to lattices of arbitrary determinant. It follows from Lemma 14. 4 that if $L \cong K$ (mod π^{2e+1}), then $L \cong K$; this is also true for 1-dimensional lattices. Hence we obtain by induction the following result of Durfee [3, Theorem 2]:

THEOREM 15. 1. *If L and K are lattices in the same space and of least divisor* $|\,\pi^m\,|$, *then* $L \cong K$ *if* $L \cong K$ (mod π^{2e+m+1}).

16. Four-dimensional lattices.

Having completed the two-dimensional case, we now concentrate on the four-dimensional improper case. We continue to use the same notation for ν. First some lemmas:

LEMMA 16. 1. *If X and Y are orthogonal vectors in the integral lattice L and if $X^2 = [\pi^p]$, $Y^2 = [\pi^q]$, $q + p$ being odd and $q < p$, then there is a unit α and a $\beta \varepsilon (\pi^{\frac{1}{2}(p-q+1)}) \cdot \mathfrak{o}$ such that $(\alpha X + \beta Y)^2 = \pi^p + \{\pi^{2e+q}\}$.*

Proof. By Lemma 4. 1 there is an α_0 such that $(\alpha_0 X)^2 = \pi^p + \{\pi^{p+1}\}$. Suppose that we have found α and β such that $(\alpha X + \beta Y)^2 = \pi^p + \epsilon\pi^r$ with the greatest possible r: to prove that $r \geqq 2e + q$. Let us assume that $r < 2e + q$; we will establish a contradiction by producing a new α and β with a bigger r. Case 1: $r + p$ is even: By Lemma 4. 1 there is a unit A such that $A^2\pi^{r-p}X^2 \equiv \epsilon\pi^r \pmod{\pi^{r+1}}$. Then $((\alpha + A\pi^{\frac{1}{2}(r-p)}) \cdot X + \beta Y)^2 = \pi^p + \{\pi^{r+1}\}$, by choice of A and since $| 2\alpha A\pi^{\frac{1}{2}(r-p)} \cdot X^2 | < | \pi^r |$, r being less than $p + 2e$. This denies the maximality of r. Case 2: $r + q$ is even. This follows in the same way as Case 1. Q. E. D.

LEMMA 16. 2. *Under the same assumptions as Lemma 16. 1, there exists a unit β and an integer α such that $(\alpha X + \beta Y)^2 = a_1\pi^q + \{\pi^{2e+q}\}$, provided that $Y^2 = a_1\pi^q + \{\pi^p\}$ where a_1 is a unit.*

Proof. Similar to proof of Lemma 16. 1.

A special change of basis. Let $L = (\mathfrak{o}X_1 + \mathfrak{o}X_2) \oplus (\mathfrak{o}X_3 + \mathfrak{o}X_4)$ be an improper totally integral lattice of unit determinant. Consider the following changes of basis in which $| \epsilon | = 1$ and $\alpha \varepsilon \mathfrak{o}$:

$$L = \mathfrak{o}X_1 + \mathfrak{o}X_2 + \mathfrak{o}(\epsilon X_3 + \alpha X_1) + \mathfrak{o}X_4 = (\mathfrak{o}Y_1 + \mathfrak{o}Y_2) \oplus (\mathfrak{o}Y_3 + \mathfrak{o}Y_4)$$

where $Y_1 = X_1 + AY_3 + BY_4$; $Y_2 = X_2 + A'Y_3 + B'Y_4$; $Y_3 = \epsilon X_3 + \alpha X_1$; $Y_4 = X_4$, the (integers) A, A', B, B', being determined by Lemma 2. 2. We denote this change of basis from (X_i) to (Y_i) by the symbol $\mathrm{op}(X_3 \rightarrow \epsilon X_3 + \alpha X_1)$ and write $Y_i = \mathrm{op}(X_i)$; this symbol is defined for any unit ϵ and for any integer α; in fact we can make preliminary estimates on the magnitudes of the A's and B's by solving the equations of Lemma 2. 2 to obtain

$$\begin{aligned} | A | &= | \alpha \cdot X_1^2 \cdot X_4^2 | \quad ; \quad | B | = | \alpha \cdot X_1^2 | \\ | A' | &= | \alpha \cdot X_4^2 \cdot | \quad ; \quad | B' | = | \alpha |. \end{aligned} \tag{16. 1}$$

Observe that $\mathrm{op}(X_1) \cdot \mathrm{op}(X_2)$ is a unit.

Definition 16. 1. A two-dimensional lattice is a null lattice if it has a basis $L = \mathfrak{o}\xi_1 + \mathfrak{o}\xi_2$ for which $\xi_1^2 = 0 = \xi_2^2$ and $\xi_1 \cdot \xi_2 = 1$.

Notation. If L is any totally integral lattice of unit determinant, $\dim L \geqq 3$, $N(L) = (\pi^\nu) \cdot \mathfrak{o}$, $0 \leqq \nu \leqq e$, we define $\mu = \min \operatorname{ord} X^2$, the minimum being over all X in L having the property $\operatorname{ord} X^2 + \nu$ is odd, provided that the μ obtained exists and is less than e; otherwise take $\mu = e$. When there is any risk of misunderstanding we will write $\mu(L)$ for μ. Note that if $\mu + \nu$ is even, then μ must be e.

LEMMA 16. 3. *There is a decomposition $L = L_0 \oplus L_1$ for L such that*

$$L_0 \cong \begin{pmatrix} [\pi^\nu] & 1 \\ 1 & \{\pi^\gamma\} \end{pmatrix} \text{ and } L_1 \cong \begin{pmatrix} \pi^\mu & 1 \\ 1 & \{\pi^{2e-\mu}\} \end{pmatrix} \tag{16. 2}$$

in which (1) *L_1 is null and $\gamma \geqq e+1$ when $\mu + \nu$ is even*; (2) *$\gamma \geqq \mu$ when $\mu + \nu$ is odd.*

Proof. There is a basis $L = (\mathfrak{o}\xi_1 + \mathfrak{o}\xi_2) \oplus (\mathfrak{o}\xi_3 + \mathfrak{o}\xi_4)$ in which (by Lemmas 4. 3, 2. 2)

$$\xi_1^2 = [\pi^\nu]\,; \quad \xi_1 \cdot \xi_2 = 1; \text{ and } \xi_2^2 \,\varepsilon\, \{\pi^\nu\}. \tag{16. 3}$$

The second component can then be made to take the form $\xi_3^2 = [\pi^\lambda]$; $\xi_3 \cdot \xi_4 = 1$; $\xi_4^2 \,\varepsilon\, (\pi^\mu) \cdot \mathfrak{o}$ in virtue of Lemma 14. 1. Of all minimal bases (ξ) satisfying the equations (16. 3) and having $\xi_4^2 \,\varepsilon\, (\pi^\mu) \cdot \mathfrak{o}$, choose one for which λ is greatest (possibly ∞). Then $\lambda \geqq \mu$. For if not, $\lambda + \nu$ would be even by definition of μ; then there would be an $\alpha_0 \,\varepsilon\, \mathfrak{o}$ such that $(\xi_3 + \alpha_0 \cdot \xi_1)^2 \,\varepsilon\, (\pi^{\lambda+1}) \cdot \mathfrak{o}$; apply $\operatorname{op}(X_3 \to X_3 + \alpha X_1)$ on this basis with $\alpha = \alpha_0$; then $(\operatorname{op}(X_1))^2 = [\pi^\nu]$ by equations (16. 1). This is a contradiction and so $\lambda \geqq \mu$.

Case 1. $\mu = e$ and $\nu + e$ is even: Then $\lambda > e$ and $(\mathfrak{o}\xi_3 + \mathfrak{o}\xi_4)$ is therefore null. Applying Lemma 14. 1 to $(\mathfrak{o}\xi_1 + \mathfrak{o}\xi_2)$ completes the proof of this case.

Case 2. $\mu + \nu$ is odd: We first show how to achieve $\lambda = \mu$ by suitable changes of basis. If $(\mathfrak{o}\xi_3 + \mathfrak{o}\xi_4)$ represents $[\pi^\mu]$ (e. g. when $\mu = e$), we are through. If not, then $(\mathfrak{o}\xi_1 + \mathfrak{o}\xi_2)$ must represent $[\pi^\mu]$ and we can therefore write $\mathfrak{o}\xi_1 + \mathfrak{o}\xi_2 = \mathfrak{o}\eta_1 + \mathfrak{o}\eta_2$ with $\eta_1^2 = [\pi^\nu]$; $\eta_1 \cdot \eta_2 = 1$; $\eta_2^2 = [\pi^\mu]$ in virtue of Lemma 14. 1. If we apply $\operatorname{op}(X_3 \to X_3 + X_2)$ to the basis $(\eta_1, \eta_2, \xi_3, \xi_4)$ we obtain

$$L \cong \begin{pmatrix} [\pi^\nu] & 1 \\ 1 & \{\pi^\nu\} \end{pmatrix} \oplus \begin{pmatrix} \delta\pi^\mu & 1 \\ 1 & [\pi^\beta] \end{pmatrix}, \qquad |\delta| = 1 \tag{16. 4}$$

in the basis $(\zeta_1, \zeta_2, \zeta_3, \zeta_4)$. Of all minimal bases (ζ) having the form of (16. 4), choose one having a maximal β. Then $\beta \geqq 2e - \mu$, for if $\beta < 2e - \mu$ then

$\beta + \mu$ would be odd by Lemma 14. 1, and an application of $\mathrm{op}(X_4 \to X_4 + \alpha X_1)$ with a suitable α would lead to a contradiction.

To show that δ can be taken equal to 1: by Lemma 16. 1 we can find a unit ϵ and an integer α such that $(\epsilon\zeta_3 + \alpha\zeta_1)^2 = \pi^{\mu} + \{\pi^{2e+\nu}\}$. Apply $\mathrm{op}(X_3 \to \epsilon X_3 + \alpha X_1)$ and call the first component L_0, the second L_1. Then L_0 has the required form by Lemma 14. 1; and L_1 has the required form by Lemma 14. 4. Q. E. D.

LEMMA 16. 4. *If L and K are lattices in the same metric space and have the forms*

$$\begin{pmatrix} a_1\pi^{\nu} & 1 \\ 1 & \{\pi^{\nu}\} \end{pmatrix} \oplus \begin{pmatrix} [\pi^{\mu}] & 1 \\ 1 & \{\pi^{2e-\mu}\} \end{pmatrix} \tag{16. 5a}$$

$$\begin{pmatrix} a_1\pi^{\nu} & 1 \\ 1 & \{\pi^{\nu}\} \end{pmatrix} \oplus \begin{pmatrix} \pi^{\mu} & 1 \\ 1 & \{\pi^{2e-\mu}\} \end{pmatrix} \tag{16. 5b}$$

respectively, and if $\mu + \nu$ is odd, then $L \cong K$.

Proof. Let L have the form (16. 5a) when written in the basis $L = (\mathfrak{o}\xi_1 + \mathfrak{o}\xi_2) \oplus (\mathfrak{o}\xi_3 + \mathfrak{o}\xi_4)$. By Lemma 16. 1 there is a unit ϵ and an $\alpha \in (\pi^{\frac{1}{2}(\mu-\nu+1)}) \cdot \mathfrak{o}$ such that $(\alpha\xi_1 + \epsilon\xi_3)^2 = \pi^{\mu} + \{\pi^{2e+\nu}\}$; apply $\mathrm{op}(X_3 \to \epsilon X_3 + \alpha X_1)$ to this basis with these values of α and ϵ. Then $(\mathrm{op}(X_3))^2 = \pi^{\mu} + \{\pi^{2e+\nu}\}$ and $(\mathrm{op}(X_1))^2 = a_1\pi^{\nu} + \{\pi^{2e}\}$ by equations (16. 1). Then applying Lemma 14. 4 to both of these new components for L shows that L can be written in the form (16. 5b). Call this decomposition $L = L_0 \oplus L_1$. Let $K = K_0 \oplus K_1$ be a decomposition of K corresponding to the form (16. 5b). In virtue of Theorem 14. 3 it suffices to prove that $d(L_0) \cong d(K_0)$ and $d(L_1) \cong d(K_1)$ and this amounts to proving that $d(L_1) \cong d(K_1)$ since $d(L) \cong d \cong d(K)$.

To prove $d(L_1) \cong d(K_1)$ we take S-symbols. Direct calculation shows that $1 \sim S(L) \cdot S(K) \sim (d(L_1) \cdot d(K_1), -a_1 d\pi)$. Utilizing the remark of Section 8 we see that this can only occur if $d(L_1) \cong d(K_1)$ since $d(L_1) \cdot d(K_1) = 1 + \{4\}$. Q. E. D.

THEOREM 16. 5. *If L and K are improper 4-dimensional lattices in the 4-dimensional metric space V and have the properties*: (1) $N(L) = (\pi^{\nu}) \cdot \mathfrak{o} = N(K)$; (2) $\mu(L) = \mu = \mu(K)$; (3) *corresponding to each Y in K with $Y^2 = [\pi^{\nu}]$, there is an X in L such that $X^2 \equiv Y^2 \pmod{\pi^{\mu}}$. Then $L \cong K$.*

Proof. We can assume that $e > \nu$. If $\mu + \nu$ is even, then $\mu = e$ and the result follows from Lemma 16. 3 and Corollary 14. 3; we can therefore assume that $\mu + \nu$ is odd. By Lemma 16. 3 we can write $K = K_1 \oplus K_0$ with

$$K_1 \cong \begin{pmatrix} \pi^{\mu} & 1 \\ 1 & \{\pi^{2e-\mu}\} \end{pmatrix} \text{ and } K_0 \cong \begin{pmatrix} a_1\pi^{\nu} & 1 \\ 1 & a_2\pi^{\lambda} \end{pmatrix} \tag{16. 6}$$

where a_1 and a_2 are units and $\lambda \geqq \mu$. Now apply Lemma 16.3 to L to obtain $L = L_1 \oplus L_0$. Then L_0 must represent $a_1\pi^\nu + \{\pi^\mu\}$ since L represents this quantity and since $N(L_1) = (\pi^\mu) \cdot \mathfrak{o}$. Hence we can find bases $L_1 = \mathfrak{o}\xi_1 + \mathfrak{o}\xi_2$ and $L_0 = \mathfrak{o}\xi_3 + \mathfrak{o}\xi_4$ in which L has the form

$$(16.7) \qquad L \cong \begin{pmatrix} [\pi^\mu] & 1 \\ 1 & \{\pi^\nu\} \end{pmatrix} \oplus \begin{pmatrix} a_1\pi^\nu + \delta\pi^\mu & 1 \\ 1 & a_2\pi^\lambda + [\pi^\Gamma] \end{pmatrix}, \quad \delta \,\varepsilon\, \mathfrak{o}$$

where $\mu \leqq \Gamma \leqq \infty$. By Lemma 16.2 there is a unit ϵ and an integer α such that $(\epsilon\xi_3 + \alpha\xi_1)^2 = a_1\pi^\nu + \{\pi^{2e+\nu}\}$. Apply $\mathrm{op}(X_3 \to \epsilon X_3 + \alpha X_1)$ to the basis (ξ): then $(\mathrm{op}(X_1))^2 = [\pi^\mu]$ and $(\mathrm{op}(X_3))^2 = a_1\pi^\nu + \{\pi^{2e+\nu}\}$ by the equations (16.1). Now apply Lemma 14.4 to the resulting second component after the application of the above operation and we see that L has a basis (ζ) in which L has the form of equation (16.7) with $\delta = 0$. Of all bases (ζ) having this form, choose one in which Γ is maximal. Then Γ must be $\geqq 2e - \nu$: for if $\Gamma < 2e - \nu$, then $\Gamma + \nu$ must be odd, using the argument of equation (14.3); in that event, $\Gamma + \mu$ would be even and an application of $\mathrm{op}(X_4 \to \epsilon X_4 + \alpha X_1)$ would lead to a contradiction. Hence $\Gamma \geqq 2e - \nu$.

Hence we have $L = L_1{}^* \oplus L_0{}^*$ with $d(L_0{}^*) \cong d(K_0) \pmod 4$ and hence $d(L_1{}^*) \cong d(K_1) \pmod 4$ and hence $d(L_1{}^*) \cong -1 \pmod 4$. Now $L_1{}^*$ represents $[\pi^\mu]$ and a determinantal consideration (compare Remark 14.1) then shows that $L_1{}^* \cong \begin{pmatrix} [\pi^\mu] & 1 \\ 1 & \{\pi^{2e-\mu}\} \end{pmatrix}$. The proof of the theorem follows by applying Lemma 16.4 to $L_0{}^* \oplus L_1{}^*$ and $K_0 \oplus K_1$.

17. The general improper case. We now consider general improper lattices L of unit determinant. Most of the labour has already been done and we need only show that the general case reduces to the case of dimension 4. By Theorem 2.4 we can write $L = \sum L_\lambda$ where the L_λ are improper and of dimension 2. Then it is possible to write $L = L_1 \oplus L_2 \oplus \sum N_\lambda$ where L_1 and L_2 are of dimension 2 and the N_λ are null lattices: this can be achieved by successively applying $\mathrm{op}(X_i \to \epsilon X_i + \alpha X_j)$ but we have manipulated with this operation often enough to be able to omit the details of the proof.

Now let L and K be lattices of arbitrary dimension but otherwise satisfying the conditions of Theorem 16.5. Reduce L and K to the forms $L_1 \oplus L_2 \oplus \sum N_\lambda$ and $K_1 \oplus K_2 \oplus \sum N_\lambda$. Then $L_1 \oplus L_2$ and $K_1 \oplus K_2$ are in isometric spaces [20] and also satisfy the remaining conditions of Theorem 16.5. Hence they are isometric. *Hence Theorem* 16.5 *holds for lattices of arbitrary dimension.*

18. The proper case. We carry over the notation for μ and ν: in this case ν will always be zero and μ can be even only when $\mu = e$. The proper case will be made to depend on the improper case. We state without proof the fact that a proper lattice $L = \mathfrak{o}\xi_1 \oplus \cdots \oplus \mathfrak{o}\xi_n$ with $|\xi_i^2| = 1$ can be written

$$L = L_1 \oplus L_2 \oplus \sum N_\lambda \tag{18.1}$$

where L_1 is proper and of dimension $\leqq 2$, the N_λ are all null and L_2 is of the form $\begin{pmatrix} \pi^\mu & 1 \\ 1 & \{\pi^{2e-\mu}\} \end{pmatrix}$ when μ is odd, and is the null lattice when μ is even: the proof is similar to the argument of Lemma 16. 3.

LEMMA 18. 1. *If L and K are proper lattices of unit determinant and of odd dimension in the same metric space V and if $\mu(L) = \mu(K)$, then $L \cong K$.*

Proof. If μ is even the proof is immediate from equation (18. 1). So assume that μ is odd. Again referring to equation (18. 1), we can write $L = \mathfrak{o}\xi_1 \oplus (\mathfrak{o}\xi_2 + \mathfrak{o}\xi_3) \oplus \sum N_\lambda$ and $K = \mathfrak{o}\eta_1 \oplus (\mathfrak{o}\eta_2 + \mathfrak{o}\eta_3) + \sum N_\lambda$. In fact we will prove that $\mathfrak{o}\xi_1 \oplus (\mathfrak{o}\xi_2 + \mathfrak{o}\xi_3)$ and $\mathfrak{o}\eta_1 \oplus (\mathfrak{o}\eta_2 + \mathfrak{o}\eta_3)$ are isometric, so let us denote these by L and K and assume that $\dim V = 3$.

Define an integer f equal to e or $e + 1$, the choice being such that f is odd. Construct the new metric space $V^* = F \cdot \zeta \oplus V$ with $\zeta^2 = \pi^{-f}$ and put $L^* = \mathfrak{o}\zeta \oplus L$ and $K^* = \mathfrak{o}\zeta \oplus K$. By Theorem 5. 1 it suffices to prove $L^* \cong K^*$.

By Lemma 4. 1 there is an integer α such that $(\xi_3 + \alpha\pi^r\zeta)^2 \in (\pi^{2e-\mu+1}) \cdot \mathfrak{o}$ where $2r = 2e - \mu + f$. Write $L^* = \mathfrak{o}\xi_1 \oplus (\mathfrak{o}\zeta + \mathfrak{o}\xi_2 + \mathfrak{o}\Xi_3)$ where $\Xi_3 = \xi_3 + \alpha\pi^r\zeta$. Then the equations of Lemma 2. 2 have integral solutions when applied to $\mathfrak{o}\xi_3 + \mathfrak{o}\Xi_3$ and we can write $L^* = \mathfrak{o}\xi_1 \oplus \mathfrak{o}Z \oplus (\mathfrak{o}\xi_2 + \mathfrak{o}\Xi_3)$ where $d(\mathfrak{o}\xi_2 + \mathfrak{o}\Xi_3) \cong -1$. Similarly $K^* = \mathfrak{o}\eta_1 \oplus \mathfrak{o}Y \oplus (\mathfrak{o}\eta_2 + \mathfrak{o}\mathrm{H}_3)$. Then $\mathfrak{o}\xi_2 + \mathfrak{o}\Xi_3 \cong \mathfrak{o}\eta_2 + \mathfrak{o}\mathrm{H}_3$ by Theorem 14. 3. Let $Z^2 = \epsilon_1\pi^{-f}$ and $\xi_1^2 = \epsilon_2$; then a determinantal consideration shows that $Y^2 = \epsilon_1\delta\pi^{-f}$ and $\eta_1^2 = \epsilon_2\delta$ where $\delta = 1 + \{4\}$. Now the S-symbols of $\mathfrak{o}\xi_1 \oplus \mathfrak{o}Z$ and $\mathfrak{o}\eta_1 \oplus \mathfrak{o}Y$ must be equal and direct computation yields $1 \sim (\delta, \pi)$ and δ must therefore be a square by Section 8. Q. E. D.

LEMMA 18. 2. *If L and K are of even dimension, then $L \cong K$ provided that* (1) $\mu(L) = \mu(K)$; (2) *L and K represent the same units.*

Proof. Since L and K are in the same metric space, there is nothing to prove when $n = 2$. So assume that $n \geqq 4$. By a slight modification to equation (18.1) we can write L in the form of equation (18.1) with L_2 of the form $\begin{pmatrix}\pi^\gamma & 1\\ 1 & \{\pi^{2e-\mu}\}\end{pmatrix}$ where $\gamma=\mu$ when μ is odd and $\gamma=e+1$ (and is therefore odd) when μ is even. Then

$$L\cong(a)\oplus(b)\oplus\begin{pmatrix}\pi^\gamma & 1\\ 1 & \{\pi^{2e-\mu}\}\end{pmatrix}\oplus\sum\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} \tag{18.2}$$

in the basis (ξ). Then of all bases (η) for K having the form of equation (18.3), choose one in which Δ is maximal:

$$K\cong(a+[\pi^\Delta])\oplus(c)\oplus\begin{pmatrix}\pi^\gamma & 1\\ 1 & \{\pi^{2e-\mu}\}\end{pmatrix}\oplus\sum\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}. \tag{18.3}$$

Contention: $\Delta=\infty$. For Δ must be $\geqq\mu$ since $\mathfrak{o}\eta_1+\mathfrak{o}\eta_2$ represent $a \pmod{\pi^\mu}$; and Δ cannot be even, for we could then write $\mathfrak{o}\eta_1\oplus\mathfrak{o}\eta_2=\mathfrak{o}(\eta_1+\alpha\eta_2)\oplus\mathfrak{o}H_2$ with α so chosen (by Lemma 4.1) that $(\eta_1+\alpha\eta_2)^2\equiv a \pmod{\pi^{\Delta+1}}$; it remains to consider Δ odd and $\geqq\gamma$; we could then write $\mathfrak{o}\eta_1\oplus(\mathfrak{o}\eta_3+\mathfrak{o}\eta_4)=\mathfrak{o}(\eta_1+\alpha\eta_3)\oplus K_1^*$ so that $(\eta_1+\alpha\eta_3)^2\equiv a \pmod{\pi^{\Delta+1}}$ by Lemmas 4.1 and 2.1 and it is easy to verify that $\mu(\mathfrak{o}\eta_3+\mathfrak{o}\eta_4)=\mu(K_1^*)$ and hence (equation 18.1) that $\mathfrak{o}\eta_2\oplus K_1^*$ could be written in a basis having the form $[1]\oplus\begin{pmatrix}\pi^\gamma & 1\\ 1 & \{\pi^{2e-\mu}\}\end{pmatrix}$. Hence Δ must be ∞ in the basis (η). Then we can apply Lemma 18.1 to $\mathfrak{o}\xi_2+\cdots+\mathfrak{o}\xi_n$ and $\mathfrak{o}\eta_2+\cdots+\mathfrak{o}\eta_n$. Hence $L\cong K$.
Q. E. D.

19. Piecing together Section 17, Lemma 18.1 and Lemma 18.2 proves

THEOREM 19.1. *If L and K are totally integral lattices of unit determinant in the same metric space V, then L is isometric to K provided that L and K represent the same numbers.*

20. Appendix. We have constantly used the invariants of [20, Satz 17]; these were shown by Witt to hold over any field in which every form of five or more variables is a zero form; e. g. over a local field with finite residue class field; compare also the result of Kaplansky. We now prove that the latter property is also true for a local field having a perfect residue class field of characteristic 2.

Proof. Let $l=\sum_1^5 a_ix_i^2$ be given in diagonal form. By considering $\pi\cdot l$ if necessary, we can assume that at least three of the a_i are units and that the remaining two (a_4 and a_5 say) are either units or primes. Construct a metric space V and a lattice L in V corresponding to l: $L = \mathfrak{o}\xi_1 \oplus \cdots \oplus \mathfrak{o}\xi_5$. If $|a_i| = 1$ all i, the result follows from equation (18.1). If $|a_4| = |\pi|$, by applying equation (18.1) to $\mathfrak{o}\xi_1 \oplus \mathfrak{o}\xi_2 \oplus \mathfrak{o}\xi_3$, we can write

$$L = \mathfrak{o}\eta_1 \oplus (\mathfrak{o}\eta_2 + \mathfrak{o}\eta_3) \oplus \mathfrak{o}\eta_4 \oplus \mathfrak{o}\eta_5 \tag{20.1}$$

with $|\eta_1^2| = 1$; $\eta_2 \cdot \eta_3 = 1$; $\eta_2^2, \eta_3^2 \in (\pi) \cdot \mathfrak{o}$; and $|\eta_4^2| = |\pi|$. Of all bases (η) having the form of equation (20.1), choose one in which $|\eta_2^2|$ is minimal. Then it follows from the fact that $\operatorname{ord} \eta_1^2 = 0$ and $\operatorname{ord} \eta_4^2 = 1$ and from Lemma 2.2, that $\eta_2^2 = 0$. Q. E. D.

Note that the proof just given (which rests on equation 18.1) depends solely on the operation $\operatorname{op}(X_i \to \epsilon X_i + \alpha X_j)$ and the fact that the residue class field is perfect and of characteristic 2.

PRINCETON UNIVERSITY, NEW JERSEY AND
OTAGO UNIVERSITY, NEW ZEALAND.

REFERENCES.

[1] E. Artin, "Algebraic numbers and algebraic functions I," *Princeton University, New York University*, 1950-1951.

[2] C. Chevalley, "L'Arithmétique dans les algèbres de matrices," *Actualités Scientifiques et Industrielles*, 323 (1936).

[3] W. H. Durfee, "Congruence of quadratic forms over valuation rings," *Duke Mathematical Journal*, vol. 11 (1944), pp. 687-697.

[4] ———, "Quadratic forms over fields with a valuation," *Bulletin of the American Mathematical Society*, vol. 54 (1948), pp. 338-351.

[5] H. Hasse, "Über die Äquivalenz quadratischer Formen im Körper der rationalen Zahlen," *Journal für die reine und angewandte Mathematik*, vol. 152 (1923), pp. 205-224.

[6] ———, "Äquivalenz quadratischer Formen in einem beliebigen algebraischen Zahlkörper," *Journal für die reine und angewandte Mathematik*, vol. 153 (1924), pp. 158-162.

[7] B. W. Jones, "An extension of a theorem of Witt," *Bulletin of the American Mathematical Society*, vol. 48 (1942), pp. 133-142.

[8] ———, "A canonical quadratic form for the ring of 2-adic integers," *Duke Mathematical Journal*, vol. 11 (1944), pp. 715-727.

[9] ———, "The Arithmetic Theory of Quadratic Forms," *Carus Monographs*, No. 10, Mathematical Association of America, 1950.

[10] ——— and W. H. Durfee, "A theorem on quadratic forms over the ring of 2-adic integers," *Bulletin of the American Mathematical Society*, vol. 55 (1949), pp. 758-762.

[11] M. Eichler, "Quadratische Formen und orthogonale Gruppen," Berlin (1952).

[12] ———, "Note zur Theorie der Kristallgitter," *Mathematische Annalen*, vol. 125 (1952), pp. 51-55.

[13] I. Kaplansky, "Quadratic Forms," *Journal of the Mathematical Society of Japan*, vol. 5 (1953), pp. 200-207.

[14] M. Kneser, "Zur Theorie der Kristallgitter," *Mathematische Annalen*, vol. 127 (1954), pp. 105-106.

[15] G. Pall, "The arithmetical invariants of quadratic forms," *Bulletin of the American Mathematical Society*, vol. 51 (1945), pp. 185-197.

[16] ———, "On the order invariants of integral quadratic forms," *Quarterly Journal of Mathematics*, vol. 6 (1935), pp. 30-51.

[17] H. Minkowski, *Gesammelte Abhandlungen* I (1911).

[18] C. Siegel, "Equivalence of quadratic forms," *American Journal of Mathematics*, vol. 63 (1941), pp. 658-680.

[19] B. L. van der Waerden, "Modern Algebra," New York (1949).

[20] E. Witt, "Theorie der quadratischen Formen in beliebigen Körpern," *Journal für die reine und angewandte Mathematik*, vol. 176 (1937), pp. 31-44.

AUTOMORPHISMS OF FINITE FACTORS.*

By I. M. Singer.

1. Introduction and notation. Let $\mathfrak{G}$ be a countable group of point measure preserving transformations on a non-atomic separable measure space $(X, \mathbf{F}, \mu)$, where X is the underlying space, $\mathbf{F}$ the σ-Boolean algebra of measurable sets and μ the measure. If $x \varepsilon X$ and $g \varepsilon \mathfrak{G}$, xg will denote the action of $\mathfrak{G}$ on X; and if E is any subset of X, $Eg = [xg; x \varepsilon E]$. Assume that (1.1) $\mu(X) = 1$. (1.2) if $E \varepsilon \mathbf{F}$, $\mu(E) \neq 0$, and g is not the identity of $\mathfrak{G}$, then there exists a subset $F \varepsilon \mathbf{F} \ni F \subset E$, $\mu(F) \neq 0$, and $\mu(F \Delta Fg) \neq 0$. (1.3) $\mathfrak{G}$ acts ergodically on $(X, \mathbf{F}, \mu)$; i.e., if $E \varepsilon \mathbf{F}$ and $\mu(E \Delta Eg) = 0$ for all $g \varepsilon \mathfrak{G}$, then $\mu(E) = 0$ or $\mu(E) = 1$. (Δ denotes the symmeric difference; this second condition means that an element of $\mathfrak{G}$, not the identity, is not the identity on any piece of X of positive measure.)

Associated with this action of $\mathfrak{G}$ on $(X, \mathbf{F}, \mu)$, F. J. Murray and J. von Neumann ([7] pp. 192-209) constructed a finite factor as follows: Let $\mathfrak{H} =$ [set of all complex valued functions ξ on $\mathfrak{G} \times X \ni$ for each $g \varepsilon \mathfrak{G}$, $\xi_g(x) = \xi(g, x) \varepsilon L_2(X, \mathbf{F}, \mu)$ and $\sum_{g \varepsilon \mathfrak{G}} \| \xi_g(x) \|_2{}^2 < \infty$, where $\| \cdot \|_2$ is the L_2 norm on $(X, \mathbf{F}, \mu)$.] It is not hard to see that $\mathfrak{H}$ is a separable Hilbert space, the inner product (ξ, η) of two functions ξ and η being $\sum_{g \varepsilon \mathfrak{G}} \int_X \xi(g, x) \bar{\eta}(g, x) d\mu$. In fact, $\mathfrak{H} = L_2(X \times \mathfrak{G})$ where the measure on $X \times G$ is the product of μ and Haar measure on $\mathfrak{G}$. Let $\mathfrak{B} = [\xi \varepsilon \mathfrak{H}; \xi(g, x) = 0$ except for a finite number of $g \varepsilon \mathfrak{G}$, and for each g, $\xi(g, x)$ is a bounded measurable function on $X]$. The following convolution-like product can be defined for $\eta \varepsilon \mathfrak{H}$ and $\xi \varepsilon B$: $\eta\xi(g, x) = \sum_{h \varepsilon \mathfrak{G}} \eta(h, x) \xi(gh^{-1}, xh^{-1})$, and $\eta\xi \varepsilon \mathfrak{H}$. If η has the property that $\| \eta\xi \| \leqq K \| \xi \|$ for all $\xi \varepsilon B$, then η can be extended to be a bounded operator L_η on $\mathfrak{H}$. The set $\mathfrak{L}$ of such L_η is a finite factor; moreover, $L_{\alpha_1\eta_1 + \alpha_2\eta_2} = \alpha_1 L_{\eta_1} + \alpha_2 L_{\eta_2}$; $L_{\eta_1\eta_2} = L_{\eta_1} L_{\eta_2}$ and $(L_\eta)^* = L_{\eta^*}$ where $\eta^*(g, x) = \bar{\eta}(g^{-1}, xg^{-1})$. Similarly, one can define the algebra of bounded right multiplications $\mathfrak{R}$ where $R_\eta \varepsilon \mathfrak{R}$ if $\| R_\eta \xi \| = \| \xi\eta \| \leqq K \| \xi \|$, $\xi \varepsilon \mathfrak{B}$. $\mathfrak{R}$ and $\mathfrak{L}$ are the commutants of one another; also the set $\mathfrak{B}_1$ of all bounded elements on the left equals the set of all bounded elements on the right, and the map

---

* Received March 9, 1954.

$L_\xi \rightarrow R_\xi$ of $\mathcal{L}$ into $\mathcal{R}$ is an anti-isomorphism of these two $*$-algebras. All these facts are proved in ([1] pp. 192-209).

From a different point of view [1], the algebraic system consisting of $\mathcal{H}$ together with multiplication of its elements (wherever the product makes sense and lies in $\mathcal{H}$) forms a simple H-system with identity. For the connections between these two viewpoints and other descriptions of this algebra see [2], [3], and [10]. Since both points of view have advantages, we will pass from one to the other freely. For us, then, $\mathcal{H}$ will denote the Hilbert space together with the above defined multiplication and adjoint.

Let $\mathcal{A} = [\xi \varepsilon \mathcal{H}; \xi(g,x) = 0, g \neq e]$. $\mathcal{A}$ is a maximal abelian subalgebra of the H-system $\mathcal{H}$ identifiable with the H-system $L_2(X, \boldsymbol{F}, \mu)$ under pointwise multiplication; $\mathcal{A} \cap \mathcal{B}_1$ can be identified with the set of essentially bounded measurable functions on $(X, \boldsymbol{F}, \mu)$. These identifications are made via the mapping: $\xi \rightarrow \xi(e,x)$. Conversely, if $a(x)$ is a bounded measurable function, let ξ^a be that element of $\mathcal{A} \ni \xi^a(e,x) = a(x)$; let $\mathcal{L}_{\mathcal{A}} = \{L_{\xi^a}\}$ and let $\mathcal{R}_{\mathcal{A}} = \{R_{\xi^a}\}$. To make the notation less cumbersome when there is little chance for confusion, we will identify ξ^a and the function $a(x)$. The restriction of the anti-isomorphism of $\mathcal{L}$ onto $\mathcal{R}$ is an isomorphism of $\mathcal{L}_{\mathcal{A}}$ onto $\mathcal{R}_{\mathcal{A}}$ since they are abelian.

In this paper, we analyze the group $\mathfrak{S}$ of those automorphisms of the H-system $\mathcal{H}$ sending $\mathcal{A}$ into $\mathcal{A}$. Since the entire H-system $\mathcal{H}$ was constructed out of the action of $\mathfrak{G}$ on X, i. e., $\mathfrak{G}$ on $\mathcal{A}$ $(\cong L_2(X))$, one would expect that a description of $\mathfrak{S}$ can be given entirely in terms of this action; in fact, we do succeed in describing $\mathfrak{S}$ in terms of the subgroup $\mathfrak{G}$ of the entire group of measure preserving transformations of $(X, \boldsymbol{F}, \mu)$. Questions about $\mathfrak{S}$ can be reduced to questions directly concerned with measure preserving transformations. We can answer some of these questions, but not others.

2. Preliminary lemmas. We devote this section to a series of computational lemmas showing what various operators on $\mathcal{H}$ look like.

Since $\mathcal{H}$ is, as a Hilbert space, isomorphic to a countable direct sum (indexed by elements of $\mathfrak{G}$) of $L_2(X, \boldsymbol{F}, \mu)$, any bounded operator T on $\mathcal{H}$ can be decomposed into an infinite matrix $\{T_{g,h}\}$, $g, h \varepsilon \mathfrak{G}$, where $T_{g,h}$ is a bounded operator on $L_2(X, \boldsymbol{F}, \mu)$, and $(T\xi)(g,x) = \sum_{h \varepsilon \mathfrak{G}} (T_{g,h}\xi_h)(x)$. (Remember that ξ_h is that element of $L_2(X, \boldsymbol{F}, \mu) \ni \xi_h(x) = \xi(h,x)$.) One can easily check that if $S \sim \{S_{g,h}\}$ and $T \sim \{T_{g,h}\}$ then

$$\alpha S + \beta T \sim \{\alpha S_{g,h} + \beta T_{g,h}\} \quad \text{and} \quad ST \sim \{\sum_{k \varepsilon \mathfrak{G}} S_{g,k} T_{k,h}\},$$

this last sum converging in the strong operator topology. We now compute the matrices corresponding to several bounded operators.

2.1 LEMMA. *If* $\xi^a \varepsilon \mathfrak{A} \cap \mathfrak{B}_1$, *then* $\xi^a\eta(g,x) = L_{\xi^a}(\eta)(g,x) = a(x)\eta(g,x)$ *and* $L_{\xi^a} \sim \{T_{g,h}\}$ *where* $T_{g,h} = 0$ *if* $g \neq h$ *and* $T_{g,g}$ *for any* $g \varepsilon \mathfrak{G}$ *is multiplication by the function* $a(x)$ *on* $L_2(X, \boldsymbol{F}, \mu)$. *Also* $\eta\xi^a(g,x) = R_{\xi^a}(\eta)(g,x) = a(xg^{-1})\eta(g,x)$ *and* $R_{\xi^a} \sim \{T_{g,h}\}$ *where* $T_{g,h} = 0$ *if* $g \neq h$ *and* $T_{g,g}$ *is multiplication by the function* $a(xg^{-1})$ *on* $L_2(X, \boldsymbol{F}, \mu)$.

Proof. The lemma follows directly from the definition of multiplication and the definition of ξ^a.

If $\tau \varepsilon \mathfrak{M}$ is a unitary element $(\tau^*\tau = \tau\tau^* = I)$, then it leads to the inner automorphism J_τ where $J_\tau(\eta) = \tau\eta\tau^*$. We are interested in those J_τ leaving $\mathfrak{A}$ invariant.

2.2 LEMMA. *Let* $E_g{}^\tau = [x \varepsilon X; \tau(g,x) \neq 0]$. *A necessary and sufficient condition that* τ *be a unitary element* $\ni \tau\mathfrak{A}\tau^* = \mathfrak{A}$ *is that the* $E_g{}^\tau$ *satisfy the following conditions:*

(i) $\mu(\bigcup_{g \varepsilon \mathfrak{G}} E_g{}^\tau) = 1$

(ii) if $g \neq h$, $\mu(E_g{}^\tau \cap E_h{}^\tau) = 0$

(iii) if $g \neq h$, $\mu(E_g{}^\tau g^{-1} \cap E_h{}^\tau h^{-1}) = 0$

(iv) $\sum_{h \varepsilon \mathfrak{G}} |\tau(h,x)|^2 = 1$ a. e. on X; and $|\tau(g,x)| = 1$, $x \varepsilon E_g{}^\tau$.

Proof. Suppose τ is unitary. Then $\tau\tau^* = I$ so that

$$\sum_{h \varepsilon \mathfrak{G}} \tau(h,x)\tau^*(gh^{-1}, xh^{-1}) = \sum_{h \varepsilon \mathfrak{G}} \tau(h,x)\tau(hg^{-1}, xg^{-1}) = 0$$

a. e. if $g \neq e$ and $\sum_{h \varepsilon \mathfrak{G}} |\tau(h,x)|^2 = 1$ a. e., proving the first part of (iv).

Since $\tau\mathfrak{A} = \mathfrak{A}\tau$, no matter what bounded measurable function $a(x)$ we choose, there exists a similar function $b(x) \ni \tau\xi^a = \xi^b\tau$. This means

$$\tau\xi^a(g,x) = a(xg^{-1})\tau(g,x) = b(x)\tau(g,x) = \xi^b\tau$$

(from 2.1); that is $a(xg^{-1}) = b(x)$ a. e. for $x \varepsilon E_g{}^\tau$. Hence $a(xg^{-1}) = a(xh^{-1})$ for almost all $x \varepsilon E_g{}^\tau \cap E_h{}^\tau$, and all bounded measurable functions $a(x)$. We now show that $\mu(E_g{}^\tau \cap E_h{}^\tau) = 0$ if $g \neq h$. For otherwise, by (1.2), there exists a set $F \subset E_g{}^\tau \cap E_h{}^\tau \ni \mu(F \Delta Fgh^{-1}) \neq 0$, so that $\mu(F \cap (Fg^{-1}h)') \neq 0$. Let $a(x)$ be the characteristic function of the set Fg^{-1}. $x \varepsilon F$ if and only if $xg^{-1} \varepsilon Fg^{-1}$ if and only if $a(xg^{-1}) \neq 0$. On the other hand $a(xh^{-1}) = 1$ if and only if $x \varepsilon Fg^{-1}h$; this means $a(xg^{-1}) = 1 \neq a(xh^{-1})$ if $x \varepsilon F \cap (Fg^{-1}h)'$. But $F \cap (Fg^{-1}h)' \subset F \subset E_g{}^\tau \cap E_h{}^\tau$, a contradiction.

Since $\sum_{h \varepsilon \mathfrak{G}} |\tau(h,x)|^2 = 1$ a. e. and (ii) holds, (i) and the second part of (iv) follow immediately.

For almost all $x \varepsilon X$, (i) and (ii) imply x lies in exactly one $E_g{}^\tau$. Then, if $g \neq e$, the first equation in this proof says $0 = \sum_h \tau(h, x)\bar{\tau}(hg^{-1}, xg^{-1})$ a. e. Thus for almost all $x \varepsilon E_k{}^\tau$, $0 = \tau(k, x)\bar{\tau}(kg^{-1}, xg^{-1})$ a. e., i. e., $\mu(E_k{}^\tau \cap E_{kg^{-1}}{}^\tau g) = 0$ if $g \neq e$. Let $h = kg^{-1}$ and let $\mu(E_k{}^\tau \cap E_h{}^\tau h^{-1}k) = 0$. Since k is measure preserving, $\mu(E_k{}^\tau k^{-1} \cap E_h{}^\tau h^{-1}) = 0$ $k \neq h$.

Suppose now that τ satisfies the conditions (i)-(iv). Condition (iv) says essentially that $\sum_{h \varepsilon \mathfrak{G}} \tau_h(x)$ defines a bounded measurable function $a(x)$ of absolute value 1 a. e. We can write $\tau = \xi^a \tau_1$, where $\tau_1(g, x) = e_g(x)$ and e_g is the characteristic function of the set $E_g{}^\tau$.

Since a is of absolute value 1, ξ^a is unitary. ξ^a leaves $\mathcal{A}$ invariant, so we need deal only with τ_1. It is not hard to show that τ_1 is unitary. We show that if $b(x)$ is bounded measurable, $\tau_1 \xi^b \tau_1{}^* \varepsilon \mathcal{A}$.

$$(\tau_1 \xi^b)\tau_1{}^*(g, x) = \sum_{h \varepsilon \mathfrak{G}} \tau_1{}^* \xi^b(h, x)\bar{\tau}_1(hg^{-1}, xg^{-1})$$
$$= \sum_{h \varepsilon \mathfrak{G}} b(xh^{-1})\tau_1(hx)\bar{\tau}(hg^{-1}, xg^{-1}) = \sum_{h \varepsilon \mathfrak{G}} b(xh^{-1}) e_h(x) e_{hg^{-1}}(xg^{-1}).$$

But if $g \neq e$, $E_h{}^\tau \cap E_{hg^{-1}}{}^\tau g$ is empty from (iii). Since $e_h(x) e_{hg^{-1}}(xg^{-1})$ is just the characteristic function of the set $E_h{}^\tau \cap E_{hg^{-1}}{}^\tau g$, we obtain $\tau \xi^b \tau^*(g, h) = 0$ a. e. if $g \neq e$ and $\tau \xi^b \tau^*(e, x) = \sum_h b(xh^{-1}) e_h(x)$.

2.3 COROLLARY. *If τ is a unitary element of $\mathcal{H} \ni \tau\mathcal{A}\tau^* = \mathcal{A}$ then $\tau = \xi^a \tau_1$ where $a(x)$ is a measurable function of absolute value 1 on X, and where $\tau_1(g, x) = e_g(x)$ the characteristic functions of sets $E_g{}^\tau$ satisfying (i)-(iii) above. This decomposition is unique. Moreover, the automorphism J_τ on the H-system $\mathcal{A}$ is the same as that induced by the following point measure preserving transformation $J_\tau{}'$ defined a. e. If $x \varepsilon E_g{}^\tau g^{-1}$, then $xJ_\tau{}' = xg$.*

Proof. All but the last statement follows from the previous proof. Suppose now $b(x)$ is the characteristic function of a set F. Then by a previous formula, $\tau \xi^b \tau^* = \xi^d$ where $d(x) = \sum_h b(xh^{-1}) e_h(x)$, i. e., $d(x)$ is the characteristic function of the set $\bigcup_h Fh \cap E_h{}^\tau$. This is exactly what $J_\tau{}'$ does to F, i. e., decomposes F into $\bigcup_h F \cap E_h{}^\tau h^{-1}$ and sends each $F \cap E_h{}^\tau h^{-1}$ into $(F \cap E_h{}^\tau h^{-1})h = Fh \cap E_h{}^\tau$.

The next lemma and corollary are contained in Lemma 2, p. 315 of Dixmier's paper [2]. We include the proofs for the sake of completeness.

2.4 LEMMA. *Let T be an operator on the Hilbert space $\mathcal{H}$ which commutes with both $\mathcal{L}_{\mathcal{A}}$ and $\mathcal{R}_{\mathcal{A}}$. Then $T \sim \{T_{g,h}\}$ where $T_{g,h} = 0$ if $g \neq h$ and each $T_{g,g}$ is multiplication by a bounded measurable function $t_g(x)$.*

Proof. If $a(x)$ is bounded measurable, $T(\xi^a\eta)(g,x) = \xi^a T(\eta)(g,x)$; $\sum_h T_{g,h}(a(x)\cdot\eta(h,x)) = a(x)\sum_h T_{g,h}\eta(h,x)$. If we choose η so that $\eta(h,x) = 0$ except when $h = h'$, and $\eta(h',x) = b(x)$, we have $T_{g,h'}(a(x)b(x)) = a(x)T_{g,h'}(b(x))$, i. e., $T_{g,h'}$ commutes with all multiplications by bounded measurable functions and hence is multiplication by one, $t_{g,h'}(x)$, itself.

Also $T(\eta\xi^a)(g,x) = T(\eta)\cdot\xi^a(g,x)$;

$$\sum_h t_{g,h}(x)a(xh^{-1})\eta(h,x) = a(xg^{-1})\sum_h t_{g,h}(x)\eta(h,x).$$

If we chose η as above,

$$t_{g,h'}(x)a(xh'^{-1})b(x) = a(xg^{-1})t_{g,h'}(x)b(x),$$

for all $a(x), b(x)$; $g, h' \varepsilon \mathfrak{G}$. Hence $t_{g,h}(x)(a(xh^{-1}) - a(xg^{-1})) = 0$. If $h \neq g$, and $F = [x; t_{g,h}(x) \neq 0]$ then for all functions $a(x)$, $a(xh^{-1}) = a(xg^{-1})$, $x \varepsilon F$. As before (in the proof of (2.1)), condition (1.2) implies $\mu(F) = 0$.

2.5 Corollary. *$\mathcal{L}_{\mathfrak{a}} \cup \mathcal{R}_{\mathfrak{a}}$, the weakly closed algebra generated by $\mathcal{L}_{\mathfrak{a}}$ and $\mathcal{R}_{\mathfrak{a}}$ is a maximal abelian sub-algebra of all bounded operators.*[1]

Proof. Lemma 2.4 shows that the commutant $(\mathcal{L}_{\mathfrak{a}} \cup \mathcal{R}_{\mathfrak{a}})'$ is abelian and hence $(\mathcal{L}_{\mathfrak{a}} \cup \mathcal{R}_{\mathfrak{a}})' \subseteqq (\mathcal{L}_{\mathfrak{a}} \cup \mathcal{R}_{\mathfrak{a}})'' = \mathcal{L}_{\mathfrak{a}} \cup \mathcal{R}_{\mathfrak{a}}$. On the other hand $\mathcal{L}_{\mathfrak{a}} \cup \mathcal{R}_{\mathfrak{a}}$ is abelian, so $\mathcal{L}_{\mathfrak{a}} \cup \mathcal{R}_{\mathfrak{a}} \subseteqq (\mathcal{L}_{\mathfrak{a}} \cup \mathcal{R}_{\mathfrak{a}})'$.

2.6 Lemma. *Let u_g, $g \varepsilon \mathfrak{G}$ be the element $\ni u_g(h,x) = 0$ if $h \neq g$, $u_g(g,x) = 1$, $x \varepsilon X$. Then $u_g u_h = u_{hg}$, $u_g^* = u_{g^{-1}}$. $\mathfrak{H}$ is generated by $\mathfrak{a}$ and $\{u_g\}$, $g \varepsilon \mathfrak{G}$. For any function, $a(x)$, $u_g\xi^a u_g^* = \xi^b$, where $b(x) = a(xg^{-1})$.*

3. The group of automorphisms leaving each element $\mathfrak{a}$ fixed.

If $S \varepsilon \mathfrak{S}$, then S restricted to $\mathfrak{a}$ is an automorphism of the abelian H-system $\mathfrak{a}$ and is induced by a measure preserving transformation S' on $(X, \boldsymbol{F}, \mu)$. If f is the characteristic function of the set F, and fS' of FS' then $S^{-1}(\xi^f) = \xi^{fS'}$. The mapping $\phi: S \rightarrow S'$ is a homomorphism of $\mathfrak{S}$ into the group $\mathfrak{M}$ of all measure preserving transformations on $(X, \boldsymbol{F}, \mu)$ for: $(TS)^{-1}(\xi^f) = S^{-1}(\xi^{fT'}) = \xi^{fT'S'}$. We denote the image of ϕ by $\mathfrak{S}'$ and the kernel of ϕ by $\mathfrak{K}$. $S \varepsilon \mathfrak{K}$ if and only if S' is the identity transformation, if and only if $S(\xi^a) = \xi^a$ for all functions $a(x)$. In this section we analyze $\mathfrak{K}$.

---

[1] More general H-systems with maximal abelian algebras $\mathfrak{a} \ni \mathcal{L}_{\mathfrak{a}} \cup \mathcal{R}_{\mathfrak{a}}$ are maximal abelian in the set of all operators have been considered by Ambrose ([2], pp. 45-46), and will be the topic of investigation with the author in a forthcoming joint paper. The procedure in Section 5 of this paper was suggested by some general results in our joint paper.

3.1 THEOREM. *$S \in \mathfrak{R}$ if and only if $S \sim \{S_{g,h}\}$, where $S_{g,h} = 0$ if $g \neq h$ and $S_{g,g}$ = multiplication by functions $s_g(x)$ satisfying*

$$\text{(i)} \quad |s_g(x)| = 1 \text{ a. e.} \qquad \text{(ii)} \quad s_{hg}(x) = s_h(xg^{-1})s_g(x) \text{ a. e.}$$

Proof. If $S \in \mathfrak{R}$, $S(\xi^a\eta) = S(\xi^a)S(\eta) = \xi^a S(\eta)$. Also $S(\eta\xi^a) = S(\eta)S(\xi^a) = S(\eta)\xi^a$ for any bounded measurable function $a(x)$. Hence $S \in (\mathcal{L}_{\mathcal{A}} \cup \mathcal{R}_{\mathcal{A}})'$. By 2.4, $\{S_{g,h}\}$ is diagonal and $S_{g,g}$ is multiplication by a function $s_g(x)$. Since S is unitary, $S^* = S^{-1}$ and therefore $S_{g,g}{}^* = S_{g,g}{}^{-1}$ implying $\bar{s}_g(x) = s_g(x)^{-1}$. This proves (i). As for (ii), $S(u_{gh}) = S(u_h u_g) = S(u_h)S(u_g)$ means $s_{gh}u_{gh} = s_h u_h s_g u_g = s_g u_h s_g u_h{}^* u_{gh}$. Since $u_h s_g u_h{}^* = \tilde{s}_g$ where $\tilde{s}_g(x) = s_g(xh^{-1})$ we get $s_{gh}(x) = s_h(x)s_g(xh^{-1})$.

Conversely, if S is a diagonal matrix of the given type satisfying (i), it is a unitary operator in $\mathfrak{H}$, leaving each element of $\mathcal{A}$ fixed. (ii) implies $S(u_g u_h) = S(u_g)S(u_h)$ and moreover $S(\xi^a u_g) = \xi^a S(u_g) = S(\xi^a)S(u_g)$. Hence S is an automorphism on $\mathfrak{B}$ and on all of $\mathfrak{H}$ because it is unitary.

3.2 THEOREM.[2] *Let $\mathfrak{J}$ be the set of all inner automorphisms of $\mathfrak{H}$. Then $\mathfrak{J} \cap \mathfrak{R} = [S \in \mathfrak{R}; s_g(x) = a(x)a(xg^{-1})$ for some $a(x)$ a measurable function of absolute value $1]$.*

Proof. If $S \in \mathfrak{J}$, $S(\eta) = \tau\eta\tau^*$ for some unitary $\tau \in \mathfrak{H}$. If $S \in \mathfrak{R}$ also, τ commutes with $\mathcal{A}$ and hence lies in $\mathcal{A}$. Thus $\tau = \xi^a$ and so $|a(x)| = 1$, $x \in X$. Also $S(u_g) = s_g u_g = \xi^a u_g \xi^{\bar{a}} = a(x)a(xg^{-1})u_g$.

3.3 THEOREM. *$\mathfrak{S}$ is a semidirect product of $\mathfrak{R}$ and a subgroup $\mathfrak{S}_1$ consisting of all $S_1 \in \mathfrak{S}$ satisfying the condition P: S_1 send u_g into $\tau_g \ni \tau_g(h, x)$ always consist of characteristic functions of sets.*

Proof. If $S \in \mathfrak{S}$,

$$\mathcal{A} = S(\mathcal{A}) = S(u_g \mathcal{A} u_{g^{-1}}) = S(u_g)S(\mathcal{A})S(u_g)^{-1} = S(u_g)\mathcal{A}S(u_g)^{-1}.$$

Also $S(u_g)$ is unitary since S preserves adjoints. Hence $S(u_g)$ is a unitary element whose corresponding inner automorphism leaves $\mathcal{A}$ invariant. By 2.3, $S(u_g) = \xi^{a_g}\tau_g$, a_g and τ_g satisfying the conditions of 2.3. Next note that

$$a_{hg}\tau_{hg} = S(u_{hg}) = S(u_g u_h) = a_g\tau_g a_h\tau_h = a_g(\tau_g a_h \tau_g{}^{-1})\tau_g\tau_h = b\tau_g\tau_h$$

---

[2] The theorems above can most neatly be described in terms of the cohomology theory of groups. For details, see [4]. Let $\mathfrak{C}$ denote the group of all measurable functions of absolute value 1 on X. The group $\mathfrak{G}$ acts on $\mathfrak{C}$ by virtue of its action on X, i.e., if $c(x) \in \mathfrak{C}$, $c_g(x) = c(xg^{-1})$. In terms of this action $Z_0 = H_0$ = circle group because $\mathfrak{G}$ acts ergodically on X and the constants are the only invariant elements $Z_1 = \mathfrak{R}$, $B_1 = \mathfrak{J} \cap \mathfrak{R}$ and $H_1 = \mathfrak{R}/\mathfrak{J} \cap \mathfrak{R}$ are the statements (3.1) and (3.2).

where $b(x) = a_g(x)\tilde{a}_h(x)$ and $\tilde{a}_h(x)$ is the image of $a_h(x)$ under the automorphism of $\mathcal{A}$ due to τ_g. The uniqueness of the decomposition in 2. 3 will show that $\tau_{hg} = \tau_g\tau_h$ if we can show that $\tau_g\tau_h(k, x)$ is the characteristic function of some set F. Suppose $\tau_g(k, x) = e_k(x)$ and $\tau_h(k, x) = f_k(x)$, where $e_h(x)$ and $f_k(x)$ are characteristic functions of the sets $E_k{}^{\tau_g}$ and $E_k{}^{\tau_h}$, the notation being that of 2. 3. Then

$$\tau_g\tau_h(k, x) = \sum_{l \varepsilon \mathfrak{G}} \tau_g(l, x)\tau_h(kl^{-1}, xl^{-1}) = \sum_l e_l(x) f_{kl^{-1}}(xl^{-1})$$

which equals the characteristic function of the set $\bigcup_{l \varepsilon \mathfrak{G}} E_l{}^{\tau_g} \cap E_{kl^{-1}}{}^{\tau_h} l = F$ because $E_l{}^{\tau_g}$ are disjoint.

Let S_1 be the following operator on $\mathfrak{H}$: $(S_1\eta)(g, x) = \sum_g S(\xi^{\eta_g})\tau_g$ where $\eta_g(x) = \eta(g, x)$. In other words $\eta = \sum_g \xi^{\eta_g} u_g$ and S_1 has the same effect on $\mathcal{A}$ as S does, but sends u_g into τ_g. We claim that S_1 is an automorphism of $\mathfrak{H}$. For, if $g \neq h$,

$$\begin{aligned} 0 = (\mathcal{A}u_g, \mathcal{A}u_h) = (S(\mathcal{A}u_g), S(\mathcal{A}u_h)) &= (S(\mathcal{A})S(u_g), S(\mathcal{A})S(u_h)) \\ &= (\mathcal{A}\xi^{a_g}\tau_g, \mathcal{A}\xi^{a_h}\tau_h) = (\mathcal{A}\tau_g, \mathcal{A}\tau_h). \end{aligned}$$

Thus the elements in the sum for $S_1(\eta)$ are orthogonal and

$$\begin{aligned} (S_1\eta, S_1\nu) = \sum_g (S(\xi^{\eta_g})\tau_g, S(\xi^{\nu_g}\tau_g) &= \sum_g (S(\xi^{\eta_g}), S(\xi^{\nu_g})) \\ &= \sum_g (\xi^{\eta_g}, \xi^{\nu_g}) = (\eta, \nu) \end{aligned}$$

i. e., S_1 is unitary.

S_1 preserves multiplication:

$$\begin{aligned} S_1(\xi^a u_g \xi^b u_h) &= S_1(\xi^a u_g \xi^b u_{g^{-1}} u_{hg}) = S(\xi^a u_g \xi^b u_{g^{-1}})\tau_{hg} \\ &= S(\xi^a)\xi^{a_g}(\tau_g S(\xi^b)\tau_{g^{-1}})\xi^{\bar{a}_g}\tau_g\tau_h = S(\xi^a)\tau_g S(\xi^b)\tau_h = S_1(\xi^a u_g)S_1(\xi^b u_h). \end{aligned}$$

Since elements of the form $\xi^a u_g$ generate $\mathfrak{H}$, and since S_1 is a unitary operator, S_1 preserves multiplication. To show S_1 preserves adjoints, it suffices to show it does so on a set of generators. It does for $\mathcal{A}$, since S does. It does on u_g, $g \varepsilon \mathfrak{G}$ for it preserves inverses.

We complete the proof by showing that $\mathfrak{S}_1$ is indeed a group. Suppose S_1 and $T_1 \varepsilon \mathfrak{S}_1$ and $S_1 u_g = \tau_g$ where $\tau_g(h, x) = e_h{}^g(x)$ the characteristic function of the set $E_h{}^g$. Similarly, $T_1 u_g = \nu_g$ where $\nu_g(h, x) = f_h{}^g(x)$, the characteristic function of the set $F_h{}^g$. Then

$$\begin{aligned} T_1S_1(u_g) = T_1(\tau_g) = T_1\Big(\sum_h e_h{}^g u_h\Big) &= \sum_h T_1(e_h{}^g)\nu_h \\ &= \sum_{k,h} T_1(e_h{}^g) f_k{}^h u_k = \sum_k \Big(\sum_h \tilde{e}_h{}^g f_k{}^h\Big) u_k \end{aligned}$$

where $T_1(e_h^g) = \tilde{e}_h^g$; and $\tilde{e}_h^g$ are characteristic functions of sets $\tilde{E}_h^g$ which are disjoint over the index h for each $g \varepsilon \mathfrak{G}$ because T_1 is an automorphism and the E_h^g enjoy this property also. We must show that $\sum_h \tilde{e}_h^g f_k^h$ is a characteristic function for each $g, k \varepsilon \mathfrak{G}$. In fact it is the characteristic function of the set $\bigcup_h (\tilde{E}_h^g \cap F_k^h)$ since the $\tilde{E}_h^g$ are disjoint for fixed $g \varepsilon \mathfrak{G}$.

3.4 Corollary. *Let $\mathfrak{I}_{\mathfrak{K}} = \mathfrak{K} \cap \mathfrak{I}$ and $\mathfrak{I}_{\mathfrak{S}_1} = \mathfrak{I} \cap \mathfrak{S}_1$. Then $\mathfrak{S}/\mathfrak{I}$ is the semidirect product of two groups $\tilde{\mathfrak{K}}$ and $\tilde{\mathfrak{S}}_1$ isomorphic respectively to $\mathfrak{K}/\mathfrak{I}_{\mathfrak{K}}$ and $\mathfrak{S}_1/\mathfrak{I}_{\mathfrak{S}_1}$.*

Proof. 2.3 shows that $\mathfrak{I}$ is the semidirect product of $\mathfrak{I}_{\mathfrak{K}}$ and $\mathfrak{I}_{\mathfrak{S}_1}$. Let $\tilde{\mathfrak{K}}$ be the image of $\mathfrak{K}$ in $\mathfrak{S}/\mathfrak{I}$ and $\tilde{\mathfrak{S}}_1$ the image of $\mathfrak{S}_1$. Then $\mathfrak{K}$ is normal and $\mathfrak{K} \cap \mathfrak{S}_1$ is the identity element of $\mathfrak{S}/\mathfrak{I}$. For if $K\mathfrak{I} = S_1\mathfrak{I}$, then there exists a $J \varepsilon \mathfrak{I} \ni J = J_1J_2^{-1}$, $J_1 \varepsilon \mathfrak{I}_{\mathfrak{K}}$ and $J_2 \varepsilon \mathfrak{I}_{\mathfrak{S}_1}$, and $KJ_1 = S_1J_2$. But $KJ_1 \varepsilon \mathfrak{S}_1$ and $S_1J_2 \varepsilon \mathfrak{S}_1$ so that $KJ_1 = I$; $K \varepsilon \mathfrak{I}$; $K\mathfrak{I} = \mathfrak{I}$. We have shown that $\mathfrak{S}/\mathfrak{I}$ is the semidirect product of $\tilde{\mathfrak{K}}$ and $\tilde{\mathfrak{S}}_1$. But $\tilde{\mathfrak{K}} = \mathfrak{K} \cdot \mathfrak{I}/\mathfrak{I} \cong \mathfrak{K}/\mathfrak{I}_{\mathfrak{K}}$ by the first isomorphism theorem; similarly $\tilde{\mathfrak{S}}_1 = \mathfrak{S} \cdot \mathfrak{I}/\mathfrak{I} \cong \mathfrak{S}_1/\mathfrak{I}_{\mathfrak{S}_1}$.

4. Some examples.

4.1 Lemma. *Suppose $\mathfrak{G}$ is a cyclic group with generator g. Let $\mathfrak{Z}$ = [the group of all measurable functions on X with absolute value 1 with pointwise multiplication] and let $\mathfrak{Y}$ = [all $a(x) \varepsilon \mathfrak{Z}$; $a(x) = b(xg^{-1})\bar{b}(x)$ for some $b(x) \varepsilon \mathfrak{Z}$]. Then $\mathfrak{K} \cong \mathfrak{Z}$ and $\mathfrak{I}_{\mathfrak{K}} \cong \mathfrak{Y}$.*

Proof. Since $\mathfrak{G}$ is ergodic, $\mathfrak{G}$ is infinite cyclic. By 3.1, $\mathfrak{K}$ is isomorphic with the set of all mappings of $\mathfrak{G}$ into $\mathfrak{Z} \ni s_{hk}(x) = s_h(xk^{-1})s_k(x)$, $h, k \varepsilon \mathfrak{G}$, the mapping given by $g \to s_g(x)$. Let $s_n(x) = s_{g^n}(x)$, $n = 0, +1, \cdots$. The map is then completely determined by $s_1(x)$ for $s_{n+1}(x) = s_n(xg^{-1})s_1(x)$ which means $s_{n+1}(x) = \prod_{m=0}^{n} s_1(xg^{-m})$. Conversely, if $s \varepsilon \mathfrak{Z}$, let $s_{n+1}(x) = \prod_{m=0}^{n} s(xg^{-m})$ and we must show that $g^n \to s_n(x)$ will satisfy $s_m(xg^{-m})s_n(x) = s_{n+m}(x)$. By substitution $s_m(xg^{-n})s_n(x) = \prod_{i=0}^{m} s(xg^{-i}g^{-n}) \prod_{j=0}^{n} s(xg^{-j}) = \prod_{i=0}^{m+n} s(xg^{-i}) = s_{n+m}(x)$. This is a typical cohomology argument. By 3.2 and the above, $\mathfrak{I}_{\mathfrak{G}} \cong \mathfrak{Y}$.

Let $(X, \mathbf{F}, \mu)$ be the measure space of a compact separable abelian group with μ Haar measure. Let g be a group element of $X \ni \{g^n\}$, $n = 0, \pm 1, \pm 2, \cdots$ is dense in X. Then right multiplication by g^{-1} is an ergodic transformation. ([6], p. 348; this paper also shows that we could use any $(X, \mathbf{F}, \mu)$ and g, providing g is an ergodic transformation with pure point spectrum.)

4.2 THEOREM. *Let $\chi_0 \varepsilon X^*$, the character group of X. If χ_0 is of infinite order, then $\chi_0 \not\varepsilon \mathfrak{Y}$, and therefore gives rise to an outer automorphism of $\mathfrak{A}$. In particular, if X is connected, $\mathfrak{R}/\mathfrak{I}$ contains a subgroup isomorphic to X^*.*

Proof. Suppose there exists $b(x) \varepsilon Z \ni b(x)\chi_0(x) = b(xg^{-1})$. Let b have the Fourier series L_2 expansion $\sum_{\chi \varepsilon X} \alpha_\chi \chi$. The previous equality shows $\sum_\chi \alpha_\chi \chi\chi_0 = \sum_\chi \alpha_\chi \bar{\chi}(g)\chi = \sum_\chi \alpha_{\chi\chi_0} \bar{\chi}\bar{\chi}_0(g)\chi\chi_0$. By the uniqueness of the expansion, $\alpha_\chi = \alpha_{\chi\chi_0}\chi\chi_0(g)$ and $|\alpha_\chi| = |\alpha_{\chi\chi_0}| = |\alpha_{\chi\chi_0^2}| = \cdots = |\alpha_{\chi\chi_0^n}|$. If χ_0 is infinite cyclic, this last shows that $\sum_\chi |\alpha_\chi|^2 = \infty$ unless all $\alpha_\chi = 0$. Since b is bounded measurable and not the 0 function, we have a contradiction. Hence $\chi_0 \not\varepsilon \mathfrak{Y}$.

If X is connected, no $\chi_0 \varepsilon X^*$, but the identity has finite order for the range would then be disconnected. Since multiplication of characters is the same as pointwise multiplication, $X^* \subseteqq \mathfrak{Z}$ and $X^* \cap \mathfrak{Y} =$ identity. Hence X is naturally isomorphic to a subgroup of $\mathfrak{Z}/\mathfrak{Y} \cong \mathfrak{R}/\mathfrak{I}_{\mathfrak{R}}$. The example just given exhibits lots of outer automorphisms.

4.3 THEOREM. *Let $X =$ circle group, $\mu =$ Haar measure, g an irrational rotation. Then $a(x) = (e^{ix} - p)/(1 - e^{ix}\bar{p})$ is in $\mathfrak{Z}$ but not in $\mathfrak{Y}$, for all $|p| < 1$.*

Proof. Since $|e^{ix} - p/1 - e^{ix}\bar{p}| = |e^{ix}|\,|1 - e^{-ix}p|/|1 - e^{ix}\bar{p}| = 1$, $a(x) \varepsilon \mathfrak{Z}$. Suppose there exists a function $b(x) \ni a(x)b(x) = b(x - g)$. Then $b(x)(e^{ix} - p) = (1 - e^{ix}\bar{p})b(x - g)$. If $b(x) \sim \sum_n \alpha_n e^{inx}$ we get

$$\sum_n (\alpha_n e^{i(n+1)x} + \bar{p}\alpha_n e^{-ing}e^{inx}) = \sum_n (\alpha_n e^{-ing}e^{inx} + p\alpha_n e^{inx})$$

or

$$\sum_n \alpha_{n-1}(1 + \bar{p}e^{-i(n-1)g})e^{inx} = \sum_n \alpha_n(e^{-ing} + p)e^{inx}.$$

By the uniqueness of the expansion,

$$|\alpha_n| = |\alpha_{n-1}|\,|1 + \bar{p}e^{-i(n-1)g}|/|p + e^{-ing}| = |\alpha_{n-1}|\,|w^{n-1} + p|/|w^n + p|$$

where $w = e^{-ig}$. An easy computation shows this recursion formula gives $|\alpha_{n+k}| = |\alpha_{n-1}|\,|w^{n-1} + p|/|w^{n+k} + p|$. Since the factor

$$|w^{n-1} + p|/|w^{n+k} + p| > (1 - |p|)/2 > 0,$$

$$\sum_{k=-1}^{\infty} |\alpha_{n+k}|^2 \geqq (1 - |p|)^2/2^2 \sum_{k=-1}^{\infty} |\alpha_{n-1}|^2 = \infty$$

so that, as before $b(x) = 0$, a contradiction. (By analogous means, one can exhibit many functions in $\mathfrak{Z}$ not in $\mathfrak{Y}$. However, the author has not succeeded in describing $\mathfrak{Z}/\mathfrak{Y}$ or those functions in $\mathfrak{Z}$ but not in $\mathfrak{Y}$ in any reasonable fashion.)

Our next example will be used for another purpose. It is well known [9] that, for the factor of all bounded operators, the double commutator of any weakly closed self-adjoint subalgebra is the algebra itself. In [5], it is shown that for any finite factor, there always exists a subfactor whose double relative commutator (relative to the factor) is larger than the subfactor. R. V. Kadison conjectured that perhaps a Galois-like theorem was true. Namely, if $\mathcal{M}_1$ is a subfactor of $\mathcal{M}$, and $\mathfrak{S}_{\mathcal{M}_1}$ the group of all automorphisms leaving each element of $\mathcal{M}_1$ fixed, then the set of all elements left fixed by $\mathfrak{S}_{\mathcal{M}_1}$ is exactly $\mathcal{M}_1$. (The failure of the double commutator theorem amounts to saying that using inner automorphisms would not suffice.) The author conjectured that perhaps some "reasonable" topology exists, rather than the weak one, for which the theorem would hold. Neither is the case as we now show.

4. 5 LEMMA. *Let* $\mathfrak{G}_1$ *be a subgroup of* $\mathfrak{G}$, *and let* $\mathcal{M}_1 = [\eta \in \mathcal{M}; \eta(g, x) = 0 \text{ if } g \notin \mathfrak{G}_1]$. $\mathcal{M}_1$ *is a weakly closed subalgebra of* $\mathcal{M}$ *containing* $\mathcal{A}$ *and is a factor if and only if* $\mathfrak{G}_1$ *is ergodic on* X. *Let* $\mathfrak{K}_1 = [S \in \mathfrak{K}; S(\eta) = \eta$ *for all* $\eta \in \mathcal{M}_1]$. *Then* $S \sim \{s_g(x)\} \in \mathfrak{K}_1$ *if and only if* $s_g(x) = 1$ a. e. *for* all $g \in \mathfrak{G}_1$.

Proof. The map $\eta \to \eta(g, x) = \eta_g(x)$ is a continuous mapping from $\mathcal{M}$ to $\mathcal{L}_2(X, \boldsymbol{F}, \mu)$ and hence $\mathcal{M}_1$ is strongly and weakly closed. The other statements are easily verified.

4. 6 THEOREM. *If* $\mathfrak{G}_1$ *is a normal subgroup of* $\mathfrak{G}$ *such that* $\mathfrak{G}/\mathfrak{G}_1$ *is not abelian, and if* $\mathfrak{G}_1$ *acts ergodically on* X, *then the set of all elements in* $\mathcal{M}$ *left fixed by* $\mathfrak{K}_1$ *is strictly larger than* $\mathcal{M}_1$. *In particular, if* $\mathfrak{G}/\mathfrak{G}_1$ *equals its commutator,* $\mathfrak{K}_1$ *consists of the identity alone.*

Proof. We first show that $\mathfrak{K}_1$, as a group, is isomorphic to $(\mathfrak{G}/\mathfrak{G}_1)^*$, the homomorphisms of $\mathfrak{G}/\mathfrak{G}_1$ into the circle group. If $S \sim \{s_g(x)\} \in \mathfrak{K}_1$, then it gives rise to the following mapping χ_S of $\mathfrak{G}/\mathfrak{G}_1$ into $\mathfrak{Z}$. For any coset $\mathfrak{G}_1 g$, we let $\chi_S(\mathfrak{G}_1 g) = s_g(x)$. This is independent of the representative of the coset since $s_{g_1 g}(x) = s_{g_1}(xg^{-1}) s_g(x) = s_g(x)$ because $s_{g_1}(x) = 1$ a. e., $g_1 \in \mathfrak{G}_1$. Since $\mathfrak{G}_1$ is normal, $s_g(x) = s_{g_1 g}(x) = s_{g g_1}(x) = s_g(x g_1^{-1}) s_{g_1}(x) = s_g(x g_1^{-1})$ i. e., $s_g(x)$ is invariant under all $g_1 \in \mathfrak{G}_1$. But $\mathfrak{G}_1$ is ergodic so that $s_g(x)$ is a scalar of absolute value 1. Hence χ_S maps $\mathfrak{G}/\mathfrak{G}_1$ into the circle group. Finally it is a homomorphism; for $\chi_S(\mathfrak{G}_1 g \cdot \mathfrak{G}_1 h) = \chi_S(\mathfrak{G}_1 gh) = s_{gh}(x)$

$= s_g(xh^{-1})s_h(x) = s_g(x)s_h(x) = \chi_S(\mathfrak{G}_1 g)\chi_S(\mathfrak{G}_1 h)$. Hence χ_S lies in $(\mathfrak{G}/\mathfrak{G}_1)^*$. If χ_S is the identity element, then $s_g(x) = 1$ a. e. and $S = I$. The map $\chi : S \to \chi_S$ is a homomorphism because $ST \sim \{s_g(x)t_g(x)\}$. χ is onto for, given any character χ_0, define $S_0 = s_g{}^0(x)$ where $s_g{}^0(x) = \chi_0(\mathfrak{G}_1 g)$ and then $\chi(S_0)$ will equal χ_0. (This is a relative cohomology argument.)

Since the circle group is abelian, $(\mathfrak{G}/\mathfrak{G}_1)^* \cong (\mathfrak{G}/\mathfrak{G}_1/[\ ])^*$ where $[\ ]$ denotes the commutator subgroup of $\mathfrak{G}/\mathfrak{G}_1$. If $[\ ] = \mathfrak{G}/\mathfrak{G}_1$, then $(\mathfrak{G}/\mathfrak{G}_1)^* = (e)$ and $\mathfrak{K}_1$ consists of the identity alone. More generally, if $\mathfrak{G}/\mathfrak{G}_1$ is not abelian, it has a commutator. Suppose $g \varepsilon \mathfrak{G}$ is an element $\ni \mathfrak{G}_1 g$ is a commutator and $\mathfrak{G}_1 g \neq \mathfrak{G}_1$. If $s \varepsilon \mathfrak{K}_1$, then $\chi_S(\mathfrak{G}_1 g) = 1 = s_g(x)$. Therefore $S(u_g) = \xi^{s_g} u_g = u_g$, i. e. u_g is left fixed by S. But this $g \not\varepsilon \mathfrak{G}_1$ and hence $u_g \not\varepsilon \mathfrak{A}_1$.

We add that if by a "reasonable topology" we mean one such that the projection map $\eta \to \eta_g(x)$ is continuous from $\mathfrak{A}$ to $\mathcal{L}_2(X, \boldsymbol{F}, \mu)$, then $\mathfrak{A}_1$ is closed in this topology and still does not satisfy the Galois theorem.

For a specific example of the situation in 4. 6, consider the measure space $(X, \boldsymbol{F}, \mu)$ where X is the circle again and μ Haar measure. $\mathfrak{G}$ is the group generated by two irrational rotations α and β not rationally connected, together with the reflection $x \to -x$. $\mathfrak{G}$ is isomorphic to the set of triples $\{(n\alpha, m\beta, \pm 1)\}$, $n, m = 0, \pm 1, \pm 2, \cdots$ in which multiplication is given by

$$(n\alpha, m\beta, -1)(n_1\alpha, m_1\beta, \pm 1) = ((n - n_1)\alpha, (m - m_1)\beta, \mp 1);$$

$$(n\alpha, m\beta, +1)(n_1\alpha, m_1\beta, \pm 1) = ((n + n_1)\alpha, (m + m_1)\beta, \pm 1).$$

The isomorphism can be set up by the condition that the triple $(n\alpha, m\beta, -1)$ acts on x by $x(n\alpha, m\beta, -1) = -x - n\alpha - m\beta$. One can now check the conditions 1. 2 and 1. 3, to be assured that we do get a finite factor. Let $\mathfrak{G}_1 = \{(n\alpha, 0, +1)\}$, $n = 0, \pm 1, \cdots$. $\mathfrak{G}_1$ is normal and moreover, $\mathfrak{G}/\mathfrak{G}_1 \cong \{(0, m\beta, \pm 1)\}$, $m = 0, \pm 1, \pm 2, \cdots$, which is not abelian.

Finally, is Section 5, the normalizer $\mathfrak{N}$ of $\mathfrak{G}$, as a subgroup of all measure preserving transformations takes on importance and we want to compute $\mathfrak{N}$ in the special case where $\mathfrak{G} = \{g^n\}$ $n = 0, \pm 1, \pm 2, \cdots$, and g is an irrational rotation on the circle. If $N \varepsilon \mathfrak{N}$ and $NgN^{-1} = g$ then by ([6], Cor. 2, p. 347), N is a rotation. If $NgN^{-1} = g^{-1}$ then $NRg = gNR$ where R is the reflection $x \to -x$. Thus $\mathfrak{N}$ is the semidirect product of the group of all rotations with the group of order two with generator $\mathfrak{R}$.

5. The subgroup $\mathfrak{S}_1$. Since $\mathfrak{S}$ is a semidirect product of $\mathfrak{K}$ and $\mathfrak{S}_1$, and since $\mathfrak{S}/\mathfrak{K} \cong \mathfrak{S}_1{}'$ a subgroup of all measure preserving transformations, $\mathfrak{S}_1 \cong \mathfrak{S}_1{}'$ and we can view $\mathfrak{S}_1$ as a group of measure preserving transformations (m.p.t.). In this section we analyze this group, in particular we study its relation to the original group $\mathfrak{G}$.

5.1 THEOREM. *A necessary and sufficient condition that a m.p.t.* S' *stem from an automorphism* $S \varepsilon \mathfrak{S}_1$ *is that there exist sets* $E_h^g \varepsilon \boldsymbol{F}, g, h \varepsilon \mathfrak{G}$ *such that*

(i) $\mu(E_h^g \cap E_k^g) = 0$ *for* $h \neq k$.

(ii) $\mu(\bigcup_h E_h^g) = 1$.

(iii) *On* $E_h^g S'$, $(S')^{-1}h^{-1}S' = g^{-1}$.

Proof. This theorem says that the m.p.t. S' will lead to an automorphism if and only if for each $g \varepsilon \mathfrak{G}$, X can be decomposed into disjoint sets $E_h^g S'$ on which $(S')^{-1}h^{-1}S'$ looks like g^{-1}.

Now if $S \varepsilon \mathfrak{S}_1$, S induces a measure preserving transformation S' on $X \ni S^{-1}(\xi^e) = \xi^f$ where e and f are the characteristic functions of the sets E and ES' respectively. As briefly described at the beginning of Section 3, S' induces an automorphism on the H system $L_2(X, \boldsymbol{F}, \mu)$ which sends $a \rightarrow a_{S'}$, and which agrees with the restriction of S to $\mathcal{A}$. So if $a(x) \varepsilon \mathcal{A}$, $S(a) = a_{S'}$. (If S' were a point transformation, $a_{S'}(x) = a(xS')$.)

As before let $S(u_g) = \tau_g$ so that $\tau_g(h, x) = e_h^g$, the characteristic functions of the sets E_h^g. By 2.2, the E_h^g satisfy (i). We now show they satisfy (ii). Let v be a unitary in $\mathcal{H} \ni v\mathcal{A}v^{-1} = \mathcal{A}$. Let $J(v)$ be the inner automorphism on $\mathcal{H}$ due to v. Then $J(v) \varepsilon \mathfrak{S}$. It is easy to check that $J(u_g) \varepsilon \mathfrak{S}_1$ and $(J(u_g))' = g^{-1}$. Since

$$(SJ(u_g))(\mathcal{A}) = S(u_g \mathcal{A} u_g^*) = \tau_g S(\mathcal{A}) \tau_g^* = (J(\tau_g)S)(\mathcal{A}),$$

we see that $S'g^{-1}(S')^{-1} = (J(\tau_g))'$. By 2.3, $(J(\tau_g))'$ looks like this: If $x \varepsilon E_h^g h^{-1}$ then $x(J(\tau_g))' = xh$, i.e., $E_h^g h^{-1} S' g^{-1} (S')^{-1} = E_h^g h^{-1} (J(\tau_g))' = E_h^g$, pointwise which was to be proved.

Conversely, suppose S' satisfies the conditions of the theorem, relative to the sets E_h^g. S', by virtue of being m.p.t., is already an automorphism of $\mathcal{A}$ by $S(\xi^a) = \xi^{a_{s'}}$. We extend S as follows: $S(u_g) = \tau_g$, where $\tau_g(h, x) = e_h^g(x)$; $S(\xi^a u_g) = S(\xi^a)\tau_g$ and now extend S to be linear on finite linear combinations of elements of the form $\xi^a u_g$. We show S is an automorphism. First, $\tau_g \tau_h = \tau_{hg}$. For $\tau_g \tau_h(k, x) = \sum_l \tau_g(l, x) \tau_h\,(kl^{-1}, xl^{-1}) = \sum_{l_1} \tau_g(l_1^{-1}k, x) \tau_h(l_1, xk^{-1}l_1) = f_k(x)$ where $F_k = \bigcup_l E_{l^{-1}k}^g \cap E_l^h l^{-1} k$ since $E_{l^{-1}k}^g$ are disjoint for distinct l. If $x \varepsilon F_k$, then for some l, $x \varepsilon E_{l^{-1}k}^g$ and $xk^{-1}l \varepsilon E_l^h$ so that $xk^{-1}lS' = xS'g^{-1}$ and $xk^{-1}ll^{-1}S' = xk^{-1}lS'h^{-1}$. Therefore $xk^{-1}S' = xS'g^{-1}h^{-1}$ and $x \varepsilon E_k^{hg}$. So far we have shown that $F_k \subseteqq E_k^{hg}$. On the other hand, by changing variables back again $F_k = \bigcup_l (E_l^g \cap E_{kl^{-1}}^{h} l)$ and $\bigcup_k F_k = \bigcup_l (E_l^g \cap (\bigcup_k E_{kl^{-1}}^{h}) l) = X$. Hence

$$0 = \mu(X - \bigcup_k F_k) = \mu(\bigcup_k E_k^{hg} - \bigcup_k F_k)$$
$$= \mu(\bigcup_k (E_k^{hg} - F_k)) = \sum_k \mu(E_k^{hg} - F_k)$$

and $f_k = e_k^{hg}$ a. e. This proves

$$\tau_g \tau_h = \tau_{hg} \text{ and } S(u_g u_h) = \tau_{hg} = \tau_g \tau_h = S(u_g) S(u_h).$$

By 2. 2, the τ_g are unitary and $(J(\tau_g))(\mathfrak{A}) \subseteqq \mathfrak{A}$ since $E_{h_1}{}^g h_1{}^{-1} \cap E_{h_2}{}^g h_2{}^{-1}$ is empty if $h_1 \neq h_2$; that is, by (iii), on $E_{h_i}{}^g h_i{}^{-1}$, $h_i S' g (S')^{-1}$ acts like the identity, $(i = 1, 2)$ so that on the intersection h_1 acts like h_2.

We now show that $S(\xi^a u_g \xi^b u_h) = S(\xi^a u_g) S(\xi^b u_h)$. Since

$$S(\xi^a u_g \xi^b u_h) = S(\xi^a (u_g \xi^b u_g{}^*) u_{hg})$$
$$= S(\xi^a) S(u_g \xi^b u_g{}^*) \tau_g \tau_h = S(\xi^a) \tau_g \tau_g{}^{-1} S(u_g \xi^b u_g{}^{-1}) \tau_g \tau_h$$

it suffices to show that $S(u_g \xi^b u_g{}^*) = \tau_g S(\xi^b) \tau_g{}^{-1}$. But $S(u_g \xi^b u_g{}^*) = \xi^c$ where $c(x) = b(xg(S')^{-1})$, while $\tau_g S(\xi^b) \tau_g{}^* = \xi^d$ where $d(x) = b(xg(S')^{-1}(J(\tau_g))')$. We must show that the m.p.t. $S'g$ equals $(J(\tau_g))'S'$. But we have already seen in 2. 3, that if $x \in E_h{}^g h^{-1}$, then $x(J(\tau_g))' = xh$. Condition (iii) says that if $x \in E_h{}^g h^{-1}$, then $xS' = xhS'g^{-1} = x(J(\tau_g))'S'g^{-1}$. Hence on $E_h{}^g h^{-1}$, $S'g = (J(\tau_g))'S'$. Since $\bigcup_h (E_h{}^g h^{-1}) = X$, $S'g = (J(\tau_g))'S'$ a. e., which was to be proved.

By linearity, then, we have shown that S preserves addition and multiplication on a dense subset of S. It also preserves adjoints since $S(u_g{}^*) = S(u_{g^{-1}}) = \tau_{g^{-1}} = (\tau_g)^* = (S(u_g))^*$ and it preserves complex conjugate on $L_2(X, \boldsymbol{F}, \mu) = \mathfrak{A}$. The extension to all of $\mathfrak{H}$ can be made as soon as we show that S preserves inner products on linear combinations of $\xi^a u_g$. A direct computation shows that it suffices to prove that $(\mathfrak{A}\tau_g, \tau_h) = 0$ if $g \neq h$. But

$$(\xi^a \tau_g, \tau_h) = \sum_k (a(x) e_k{}^g, e_k{}^h) = \sum_k \int_{E_k{}^g \cap E_k{}^h} a(x)\, d\mu = 0$$

if $\mu(E_k{}^g \cap E_k{}^h) = 0$ when $g \neq h$. By (iii), on $E_k{}^g S'$, $(S')^{-1} k^{-1} S' = g^{-1}$ and on $E_k{}^h S'$, $(S')^{-1} k^{-1} S' = h^{-1}$ so that $g^{-1} = h^{-1}$ on

$$E_k{}^g S' \cap E_k{}^h S' = (E_k{}^g \cap E_k{}^h) S'.$$

Hence $\mu(E_k{}^g \cap E_k{}^h) = 0$.

The theorem just proved has an interesting geometric interpretation,[1] which we are about to describe. No proofs will be given, since the proofs would amount to a reshuffling of the statements in the proof of 5. 1. We introduce a measure m on the product space $(X \times X, \boldsymbol{F} \times \boldsymbol{F})$ which will be concentrated on the graphs of g, $g \varepsilon \mathfrak{G}$, i. e., concentrated on the sets of the form $X^g = [(x, xg^{-1}), x \varepsilon X]$. For any $E \varepsilon \boldsymbol{F}$ let $E^g = [(x, xg^{-1}), x \varepsilon E]$. Then the measure m will be completely determined by the conditions that m be concentrated on X^g and $m(E^g) = \mu(E)$, i. e., the map $x \rightarrow (x, xg^{-1})$ is measure preserving into, for all $g \varepsilon \mathfrak{G}$.

There is a natural isometry of the Hilbert space $\mathfrak{H}$ onto $L_2(X \times X, \boldsymbol{F} \times \boldsymbol{F}, m)$ which is given by sending any $\eta \varepsilon \mathfrak{H}$ into the function η' on $X \times X$ where $\eta'(x, y) = 0$ if $y \neq xg^{-1}$ for some $g \varepsilon \mathfrak{G}$; and $\eta'(x, xg^{-1}) = \eta(g, x)$. (This is well defined, for by (1. 2), $(x, xg^{-1}) = (x, xh^{-1})$ only on an m-set of m-measure 0.) This isometry says little more than the fact that $\mathfrak{H}$ equals $\sum_{\mathfrak{G}} \oplus L_2(X, \boldsymbol{F}, \mu)$ as does $L_2(X \times X, \boldsymbol{F} \times \boldsymbol{F}, m)$. However, the important thing in this mapping is that multiplication in $\mathfrak{H}$ goes over to ordinary matrix multiplication in $L_2(X \times X, \boldsymbol{F} \times \boldsymbol{F}, m)$. By ordinary matrix multiplication, we mean $\eta'\tau'(x, y) = \sum_{g} \eta'(x, z)\tau'(z, y)$. We use the ordinary sum because, since $\mathfrak{G}$ is countable, there is at most a countable number of values z on which $\eta'(x, z) \neq 0$. (If $\mathfrak{G}$ were the group of order n generated by the cycle $(12 \cdots n)$ acting on the space X of n elements, the above description would lead to the algebra of all $n \times n$ matrices.)

To each m.p.t., S' on $(X, \boldsymbol{F}, \mu)$, we associate a transformation $S' \times S'$ on $(X \times X, \boldsymbol{F} \times \boldsymbol{F})$ which amounts to $(x, y)S' \times S' = (xS', yS')$ if S' were a point transformation. Then Theorem 5. 1 can be rephrased to say: *A necessary and sufficient condition a m.p.t. S' stem from an automorphism $S \varepsilon \mathfrak{S}_1$ is that $S' \times S'$ preserve the m-null sets of $(X \times X, \boldsymbol{F} \times \boldsymbol{F}, m)$. This is to be expected since the carrier of m carries all the algebraic structure of the H-system.* We leave the proof to the reader, but wish to point out that in this context,

$$(X^g)S' \times S' \cap X^h = [(y, yh^{-1})\,; y(S')^{-1}g^{-1}S' = yh^{-1}]$$
$$= [(y, yh^{-1})\,; y \varepsilon E_g{}^h S'].$$

5. 2 THEOREM. *Let $\mathfrak{N}$ denote the normalizer of $\mathfrak{G}$ in the group of all m.p.t. Then $\mathfrak{N} \subseteqq \mathfrak{S}$, and if $N' \varepsilon \mathfrak{N}$, then the associated $E_h{}^g = X$ if $h^{-1}N' = N'g^{-1}$ and $E_h{}^g$ is empty otherwise. If τ is a unitary $\ni J(\tau) \varepsilon \mathfrak{J}_{S_1}$ then $\tau(g, x) = f_g(x)$ and the corresponding*

$$E_h{}^g = \bigcup_k (F_{hk} \cap F_{kg}g^{-1})k^{-1} = \bigcup_l (F_l \cap F_{hlg^{-1}g})l^{-1}$$

where $f_g(x)$ is the characteristic function of F_g.

Proof. The first part of the theorem follows from the definition of normalizer. For the second statement, (2.3) says that if $x \varepsilon F_{g^{-1}g}$, then $x(J(\tau))' = xg^{-1}$. Suppose that $x \varepsilon F_{hk^{-1}k} \cap F_{k^{-1}g}g^{-1}k$. Then $xh^{-1} \varepsilon F_{hk^{-1}}hk^{-1}$ and $xh^{-1}(J(\tau))' = xh^{-1}hk^{-1} = xk^{-1}$; $x \varepsilon F_{k^{-1}g}(k^{-1}g)^{-1}$ and $x(J(\tau))' = xk^{-1}g$. Hence $xh^{-1}(J_\tau)' = x(J(\tau))'g^{-1}$ so that $x \varepsilon E_h{}^g$, i. e.,

$$F_h{}^g = \bigcup_k (F_{hk^{-1}} \cap F_{k^{-1}g}g^{-1})k \subseteqq E_h{}^g.$$

However,

$$\bigcup_h F_h{}^g = \bigcup_k (\bigcup_h F_{hk^{-1}} \cap F_{k^{-1}g}g^{-1})k = \bigcup_k F_{k^{-1}g}g^{-1}k = X,$$

so that $\mu(E_h{}^g - F_h{}^g) = 0$. The second form for $E_h{}^g$ can be obtained by making the change of variables from k to $l = gk$.

For a complete description of $\mathfrak{S}_1/\mathfrak{J}_{\mathfrak{S}_1}$, it would be nice to know a great deal about $\mathfrak{N} \cap \mathfrak{J}_{\mathfrak{S}_1}$ and $\mathfrak{N} \cup \mathfrak{J}_{\mathfrak{S}_1}$. We collect this information below.

5.3 LEMMA. $\mathfrak{G} \subseteqq \mathfrak{N} \cap \mathfrak{J}_{\mathfrak{S}_1}$. *Suppose* $T' = (J(\tau))' \varepsilon \mathfrak{N} \cap \mathfrak{J}_{\mathfrak{S}_1}$ *and* $\tau(g, x) = f_g(x)$ *as above. If* $\mu(F_l) > 0$, *the subgroup*

$$\mathfrak{G}_T{}^l = [g \varepsilon \mathfrak{G} ; lgl^{-1} = T'g]$$

is of finite index r, *where* $r \leqq 1/\mu(F_l)$.

Proof. $(J(u_g))' = g^{-1}$ shows $\mathfrak{G} \subseteqq \mathfrak{N} \cap \mathfrak{J}_{\mathfrak{S}_1}$. If $T' = (J(\tau))'$ then by 5.2, the corresponding $E_h{}^g = \bigcup_l (F_l \cap F_{hlg^{-1}g})l^{-1}$. If in addition, $T' \varepsilon \mathfrak{N}$ then $E_h{}^g$ is empty unless $h = (T')^{-1}gT' = g'$. This means $F_l \cap F_{hlg^{-1}g}$ is empty unless $h = g'$ and for each l, $F_l = F_{g'lg^{-1}g}$ and therefore $m(F_l) = m(F_{g'lg^{-1}})$. If $g_1, \cdots, g_s$ lie in distinct left cosets of $\mathfrak{G}_T{}^l$, then $g_i{}'lg_i \neq g_j{}'lg_j$, $i \neq j$ and $i, j = 1, \cdots, s$. Therefore the sets $F_{g_i{}'lg_i}$ are disjoint. They have the same measure and yet $\mu(X) = 1$. Hence $s \leqq 1/\mu(F_l)$.

The lemma shows that for any $g \varepsilon \mathfrak{G}$, $g^n \varepsilon \mathfrak{G}_T{}^l$ for some $n \leqq r$. In particular if G has unique n-th roots, when they exist, (for example if G is free abelian) then $lg^nl^{-1} = (T')^{-1}g^nT'$ implies $lgl^{-1} = (T')^{-1}gT'$ for all $g \varepsilon \mathfrak{G}$, i. e. $\mathfrak{N} \cap \mathfrak{J}_{\mathfrak{S}_1} = \mathfrak{G}$. This is surely the case in the example discussed in 4.1. In this case, then, $\mathfrak{S}_1/\mathfrak{J}_{\mathfrak{S}_1} \cong \mathfrak{S}'_1/\mathfrak{J}'_{\mathfrak{S}_1}$ contains $\mathfrak{N} \cup \mathfrak{J}'_{\mathfrak{S}_1}/\mathfrak{J}'_{\mathfrak{S}_1}$ as a subgroup. By the first isomorphism theorem for groups, this last group is isomorphic to $\mathfrak{N}/\mathfrak{N} \cap \mathfrak{J}'_{\mathfrak{S}_1}$ which becomes $\mathfrak{N}/\mathfrak{G}$. At the end of Section 4, $\mathfrak{N}$ is computed for a special case, and we see that $\mathfrak{N}/\mathfrak{G}$ has many elements, i. e., that there are many outer automorphisms in $\mathfrak{S}_1$.

The author has been unable to decide the quesion: Is $\mathfrak{N} \cup \mathfrak{J}'_{\mathfrak{S}_1} = \mathfrak{S}'_1$? However the following lemma is relevant to this question.

5.4 LEMMA. *If* $S' \varepsilon \mathfrak{S}'_1$ $S' \sim \{E_h{}^g\}$ *and* $T' \varepsilon \mathfrak{S}'_1$, $T' \sim \{F_h{}^g\}$ *then* $S'T' = (ST)' \sim \{\bigcup_k E_h{}^k \cap (F_k{}^g(S')^{-1})\}$.

Proof. The notation in the lemma means that on $E_h{}^gS'$ the transformation $(S')^{-1}h^{-1}S' = g^{-1}$ and on $F_h{}^gT'$ the transformation $(T')^{-1}h^{-1}T' = g^{-1}$. If $x \varepsilon E_h{}^k$ and $x \varepsilon F_k{}^g(S')^{-1}$, then $xh^{-1}S' = xS'k^{-1}$. Also $xS' \varepsilon F_k{}^g$ so that $xS'k^{-1}T' = xS'T'g^{-1}$. A combination of these two equation gives $xh^{-1}S'T' = xS'T'g^{-1}$, so that $x \varepsilon G_h{}^g$ where $S'T' = (ST)' \sim \{G_h{}^g\}$. We have shown that $\bigcup_k (E_h{}^k \cap (F_h{}^g(S')^{-1})) \subseteqq G_h{}^g$. These two sets are in fact equal; for

$$\bigcup_h \bigcup_k (E_h{}^k \cap (F_k{}^g(S')^{-1})) = \bigcup_k ((\bigcup_h E_h{}^k) \cap F_k{}^g(S')^{-1}) = \bigcup_k X \cap F_k{}^g(S')^{-1} = X.$$

If $S' = (J(\tau))' \varepsilon \mathfrak{J}'_{\mathfrak{S}_1}$ and $T' \varepsilon \mathfrak{N}$, then $F_k{}^g = X$ if $k = (T')^{-1}gT'$ and empty otherwise so that $G_h{}^g = E_h{}^k$, and by 5.2, $E_h{}^k = \bigcup_l (F_l \cap F_{hlk^{-1}k})l^{-1}$. To prove that $\mathfrak{N} \cup \mathfrak{J}'_{\mathfrak{S}_1} = \mathfrak{S}'_1$ then, one must show that given a collection of sets $\{G_h{}^g\}$ satisfying the conditions in (5.1), one can find a collection of sets $\{F_k\}$ satisfying (2.2) $\ni G_h{}^k = \bigcup_l (F_l \cap F_{hlk^{-1}k})l^{-1}$ for some $k \varepsilon G$ where k would equal $(T')^{-1}gT'$. We have been unable to prove this or find a counterexample, even in the simplest case discussed at the end of Section 4.

6. Concluding remarks. If the assumption of the ergodicity of $\mathfrak{G}$ is relaxed, the H-system obtained will no longer be simple. Nevertheless, it should be fairly clear to the expert that, with enough techniques, the constructions in this paper can still be carried through in that more general context. We have not done this; firstly, because it would obscure the basic ideas in the paper, and secondly, because the author is primarily interested in simple systems. The direct integral decomposition theory has been sufficiently well worked out so that almost all questions about general systems reduce to questions about simple ones.

However, there is a generalization of the Murray-von Neumann construction which we find more interesting. We can view the simple H-system $\mathcal{H}$ under discussion as having been obtained by taking a maximal abelian self adjoint algebra of all bounded operators on a Hilbert space (in our case $L_2(X, \boldsymbol{F}, \mu) = \mathcal{A}$) together with a countable group of unitaries leaving $\mathcal{A}$ invariant (in our case, the unitaries induced by the m.p.t.'s $g \varepsilon G$), and performing the original tensor product construction. Clearly $\mathcal{A}$ need not be so special. $\mathcal{A}$ could be any algebra, or equivalently, any H-system, and $\mathfrak{G}$ any countable group of automorphisms of $\mathcal{A}$. If enough conditions are put on the action of $\mathfrak{G}$ on $\mathcal{C}$, the tensor product construction will give simple systems, conceivably some new ones. We hope to give a full discussion of this situation, particularly as regards automorphisms, in another paper.

UNIVERSITY OF CALIFORNIA, LOS ANGELES.

BIBLIOGRAPHY.

[1] W. Ambrose, "The L_2 system of a unimodular group. I," *Transactions of the American Mathematical Society*, vol. 65 (1949), pp. 27-48.

[2] J. Dixmier, "Algèbres quasi unitaires," *Commentarii Mathematici Helvetici*, vol. 26 (1952), pp. 275-322.

[3] H. A. Dye, "The Radon-Nikodym theorem for finite rings of operators," *Transactions of the American Mathematical Society*, vol. 72 (1952), pp. 243-280.

[4] Samuel Eilenberg and Saunders MacLane, "Cohomology theory in abstract groups. I," *Annals of Mathematics*, vol. 48 (1947), pp. 51-78.

[5] Bent Fuglede and R. V. Kadison, "On a conjecture of Murray and von Neumann," *Proceedings of the National Academy of Sciences*, vol. 37 (1951), pp. 420-425.

[6] Paul R. Halmos and J. von Neumann, "Operator methods in classical mechanics. II," *Annals of Mathematics*, vol. 43 (1942), pp. 322-350.

[7] F. J. Murray and J. von Neumann, "On rings of operators," *Annals of Mathematics*, vol. 37 (1936), pp. 116-229.

[8] J. von Neumann, "On rings of operators. III," *Annals of Mathematics*, vol. 41 (1940), pp. 94-161.

[9] ———, "Zur Algebra der Functionaloperatoren," *Mathematische Annalen*, vol. 102 (1929), pp. 370-427.

[10] I. E. Segal, "A non-commutative extension of abstract integration," *Annals of Mathematics*, vol. 57 (1953), pp. 401-457.

REDUCTION OF ALGEBRAIC VARIETIES WITH RESPECT TO A DISCRETE VALUATION OF THE BASIC FIELD.*

By Goro Shimura.

In the arithmetic theory of algebraic varieties it is useful to have a general theory on "reduction modulo $\mathfrak{p}$" of varieties. Here "reduction" means, roughly spoken, the following process: let V be an algebraic variety defined over a field k with a discrete valuation, and denote by $\mathfrak{p}$ the maximal ideal of that valuation; then we have an algebraic variety $\bar{V}$ over the residue field κ of $\mathfrak{p}$ defined by the equations for V modulo $\mathfrak{p}$; such a variety $\bar{V}$ may be called the variety obtained from the variety V by the reduction modulo $\mathfrak{p}$. When one tries to construct the general theory of reduction in this sense, the first thing to do will be to investigate how the algebraico-geometric properties of V behave in the reduction process. The object of this paper is concerned with that problem.

We shall give a precise definition of the reduction of varieties and the reduction of cycles on a variety. In the case when the ambient varieties V and $\bar{V}$ are both absolutely irreducible, we shall show that the reduction of cycles on V defines a linear mapping from the group of rational V-cycles over k into the group of rational $\bar{V}$-cycles over κ of the same dimension; and shall show that the operations on cycles (the intersection-product, the direct product and the algebraic projection) are all preserved by that mapping (Section 4). If we consider reduction as specialization, this result may be regarded as a specialization-theory of cycles on algebraic varieties, in a certain sense. In fact, from this reduction-theory of cycles we can easily obtain a more general specialization-theory of cycles (Section 5). The latter will be a generalization of the specialization-theory of cycles which has been given by T. Matsusaka [4] and P. Samuel [6]. Our theory may be also called a "local" theory of varieties at a divisor $\mathfrak{p}$, whereas $\mathfrak{p}$ is not a usual divisor on varieties. We shall also show the following fact: let K be a field with a set of infinitely many valuations which satisfies a certain condition (condition (I) in Section 6); if V is an absolutely irreducible variety defined over K, then for almost all those valuations, it remains absolutely irreducible by the reduction process.

---

* Received February 15, 1954.

M. Deuring has given a reduction-theory of an algebraic function-field of one variable (Deuring [1]) and applied it to his theory of complex multiplication (Deuring [2]). Our results will supply an algebraico-geometric method for these theories.

We proceed in following the method of A. Weil in his book "Foundations of algebraic geometry."[1] We consider specializations over local domains,[2] and by extending Weil's theory to our case, we obtain our results without difficulty. As in Weil's book we shall first deal with affine varieties, and then proceed to abstract and projective varieties.

1. Specialization over an S-domain.

We shall call a commutative ring with the identity element an *S-ring* if the set of all non-units forms an ideal. An S-ring which has no zero-divisor will be called an *S-domain.* We shall call a Noetherian S-ring (or S-domain) a *local ring* (or *local domain*). If $\mathfrak{R}$ is a local ring and $\mathfrak{m}$ is the maximal ideal of $\mathfrak{R}$, then we have a topology on $\mathfrak{R}$ by the powers of $\mathfrak{m}$. With respect to this topology, $\mathfrak{R}$ has a *completion* $\mathfrak{R}^*$, which is a local ring containing $\mathfrak{R}$ as subring and subspace, and in which $\mathfrak{R}$ is dense. A valuation ring is an S-domain; and it is a local domain if it is discrete or is a field.

Let $\mathfrak{R}$ be an *S-domain* with the quotient field K and the maximal ideal $\mathfrak{m}$, and $\mathrm{K} = \mathfrak{R}/\mathfrak{m}$ its residue field. For convenience we consider two "*universal domains*" $\boldsymbol{K}$ and $\mathfrak{K}$ which are algebraically closed fields of infinite degree of transcendency over K and K, respectively. We call an element of $\boldsymbol{K}$ (or $\mathfrak{K}$) a *quantity of* $\boldsymbol{K}$ (or $\mathfrak{K}$).[3] Let $(x) = (x_1, \cdots, x_n)$ be a set of n quantities in $\boldsymbol{K}$ and $(\xi) = (\xi_1, \cdots, \xi_n)$ a set of n quantities in $\mathfrak{K}$. We say that (ξ) is a *specialization of* (x) *over* $\mathfrak{R}$ and denote $(x) \xrightarrow{\mathfrak{R}} (\xi)$ if the natural homomorphism of $\mathfrak{R}$ to $\mathfrak{R}/\mathfrak{m}$ can be extended to a homomorphism of $\mathfrak{R}[x]$ onto $\mathrm{K}[\xi]$ which maps (x) on (ξ). It is easy to see that (ξ) is a specialization of (x) over $\mathfrak{R}$ if and only if for every polynomial $F(X)$ in $\mathfrak{R}[X]$ such that $F(x) = 0$ we have $\bar{F}(\xi) = 0$ where $\bar{F}(X)$ is the class of $F(X)$ modulo $\mathfrak{m}$. As in [WF] we add an *infinite* element ∞ both to $\boldsymbol{K}$ and $\mathfrak{K}$ with operations $\infty^{-1} = 0$ and $0^{-1} = \infty$, and by a *generalized quantity* of $\boldsymbol{K}$ (or $\mathfrak{K}$) we understand either a quantity of $\boldsymbol{K}$ (or $\mathfrak{K}$) or ∞. So we say that a set of generalized quantities (x) in $\boldsymbol{K}$ (or (ξ) in $\mathfrak{K}$) is *finite* or *infinite* according as none of the x_j (or ξ_j) is ∞ or otherwise. Let (x) be a set of n generalized quantities in $\boldsymbol{K}$ and (ξ) a set of n generalized quantities in $\mathfrak{K}$; we call (ξ) a *specialization of* (x) *over* $\mathfrak{K}$, if there exists a set of n integers $\epsilon_j = \pm 1$, such that $(x') = (x_1^{\epsilon_1}, \cdots, x_n^{\epsilon_n})$ and $(\xi') = (\xi_1^{\epsilon_1}, \cdots, \xi_n^{\epsilon_n})$ are both finite and (ξ') is a specialization of (x') over $\mathfrak{R}$; and also denote by $(x) \xrightarrow{\mathfrak{R}} (\xi)$. We can easily show that if we have $(x) \xrightarrow{\mathrm{K}} (x')$ and $(x') \xrightarrow{\mathfrak{R}} (\xi)$ then we have $(x) \xrightarrow{\mathfrak{R}} (\xi)$ and that if we have $(x) \xrightarrow{\mathfrak{R}} (\xi)$ and $(\xi) \xrightarrow{\mathrm{K}} (\xi')$ then we have $(x) \xrightarrow{\mathfrak{R}} (\xi')$.

---

[1] Hereafter we shall use the same notations and terminologies as in this book. We shall quote this book as [WF].

[2] Such a specialization has been defined in D. G. Northcott [5] and applied to prove O. Zariski's main theorem on birational correspondences.

[3] The elements of K will be denoted by Latin, those of $\mathfrak{K}$ by Greek letters.

The following theorem is well-known. A proof is given in Weil [7] (Theorem 1.) for a more general case.

THEOREM 1. *Let* (x) *and* (y) *be two sets of generalized quantities in* $\boldsymbol{K}$. *Then every specialization* $(x) \xrightarrow{\mathfrak{R}} (\xi)$ *can be extended to a specialization* $(x, y) \xrightarrow{\mathfrak{R}} (\xi, \eta)$.

This theorem asserts that, for every S-domain $\mathfrak{R}$ contained in a field K_1, there exists a valuation ring $\mathfrak{R}_1$ with the quotient field K_1 such that $\mathfrak{R}_1 \supset \mathfrak{R}$, $\mathfrak{m}_1 \cap \mathfrak{R} = \mathfrak{m}$ where $\mathfrak{m}$ and $\mathfrak{m}_1$ are respectively the maximal ideals of $\mathfrak{R}$ and $\mathfrak{R}_1$.

We say that (η) is a *specialization of* (y) *over* $(x) \xrightarrow{\mathfrak{R}} (\xi)$ if (ξ, η) is a specialization of (x, y) over $\mathfrak{R}$.

Let (ξ) be a finite specialization of (x) over $\mathfrak{R}$. We call the set $\{F(x)/G(x) \mid \bar{G}(\xi) \neq 0\}$ the *specialization ring of* (ξ) *in* $K(x)$, where $F(X)$ and $G(X)$ are polynomials in $\mathfrak{R}[X]$, and $\bar{G}(X)$ is the class of $G(X)$ modulo $\mathfrak{m}$.[3'] This set is also an S-domain and will be denoted by $[(x) \xrightarrow{\mathfrak{R}} (\xi)]$. If $\mathfrak{R}$ is a local domain, the specialization ring $[(x) \xrightarrow{\mathfrak{R}} (\xi)]$ is also a local domain.

Let (x) be a set of quantities in $\boldsymbol{K}$; we say that (x) is *finite over* $\mathfrak{R}$ if every specialization (ξ) of (x) over $\mathfrak{R}$ is finite. The following proposition is also well-known (Weil [7], Theorem 2.).

PROPOSITION 1. *A set of quantities* $(x) = (x_1, \cdots, x_n)$ *is finite over* $\mathfrak{R}$ *if and only if every one of the* x_j *is integral over* $\mathfrak{R}$.

The following two propositions are easily proved.

PROPOSITION 2. *Let* (ξ) *be a specialization of* (x) *over* $\mathfrak{R}$. *If* $\mathfrak{R}$ *is a valuation ring, then* $\dim_{\mathrm{K}} (\xi) \leq \dim_K (x)$.

PROPOSITION 3. *Let* (x) *be a set of* n *independent variables over* K *and* (ξ) *a set of* n *independent variables over* K. *If* $\mathfrak{R}$ *is a* (*discrete*) *valuation ring, then* $[(x) \xrightarrow{\mathfrak{R}} (\xi)]$ *is also a* (*discrete*) *valuation ring.*

[3'] Hereafter we use the same notation for the class of a polynomial modulo the maximal ideal. Also we shall use the notation $\bar{k}$ for the algebraic closure of a field k.

LEMMA 1. *Let A be an integral domain and $(t_1, \cdots, t_n)$ a set of n independent variables over A. If A is integrally closed, then the polynomial domain $A[t]$ is also integrally closed.*

Proof. It is sufficient to prove our lemma in the case $n = 1$. It is well-known that an integral domain is integrally closed if and only if it is an intersection of valuation rings. (This follows from Theorem 1 and Proposition 1.) By this fact, we have a representation of A as an intersection $\cap \mathfrak{R}_\omega$ of valuation rings $\mathfrak{R}_\omega$. Denote by K the quotient field of A and by $\mathfrak{R}'_\omega$ the specialization ring $[t \xrightarrow{\mathfrak{R}_\omega} \tau_\omega]$ where τ_ω is an independent variable over the residue field of $\mathfrak{R}_\omega$. Then, by Proposition 3, $\mathfrak{R}'_\omega$ is also a valuation ring for every ω. It can be easily verified that

$$A[t] = \bigcap_\omega \mathfrak{R}_\omega \cap K[t].$$

As is well-known $K[t]$ is integrally closed; so that $A[t]$ is integrally closed.

PROPOSITION 4. *Let $(t_1, \cdots, t_n)$ be a set of n independent variables over K and $(\tau_1, \cdots, \tau_n)$ a set of n quantities in $\mathfrak{K}$. If $\mathfrak{R}$ is an integrally closed S-domain, then the specialization ring $[(t) \xrightarrow{\mathfrak{R}} (\tau)]$ is also integrally closed.*

This follows immediately from Lemma 1.

PROPOSITION 5. *Let (ξ) be a finite specialization of (x) over $\mathfrak{R}$ such that the specialization ring $\mathfrak{S} = [(x) \xrightarrow{\mathfrak{R}} (\xi)]$ is integrally closed. If $\mathfrak{R}$ is a (discrete) valuation ring, and if $\dim_K (x) = \dim_{\mathrm{K}} (\xi)$, then $\mathfrak{S}$ is also a (discrete) valuation ring.*

Proof. Let y be a quantity in $K(x)$ which is not in $\mathfrak{S}$. Then, by Proposition 1 and by our assumption, we have $y \xrightarrow{\mathfrak{S}} \infty$. If η is any finite specialization of y over $\mathfrak{S}$, then, by Proposition 2, we have

$$\dim_{\mathrm{K}} (\xi, \eta) \leqq \dim_K (x, y) = \dim_K (x) = \dim_{\mathrm{K}} (\xi).$$

This shows that η is algebraic over $\mathrm{K}(\xi)$. Let t be an independent variable over $K(x)$ and τ an independent variable over $\mathrm{K}(\xi)$. Put $\mathfrak{T} = [t \xrightarrow{\mathfrak{S}} \tau]$ and $z = (t + y)^{-1}$. By Proposition 4, $\mathfrak{T}$ is integrally closed; and it is easy to see that z is finite over $\mathfrak{T}$; so that z is contained in $\mathfrak{T}$, and has a uniquely determined specialization over $t \xrightarrow{\mathfrak{S}} \tau$. This specialization of z must be 0, since $y \xrightarrow{\mathfrak{S}} \infty$; this shows that y has the only specialization ∞ over $\mathfrak{S}$. Hence y^{-1} is finite over $\mathfrak{S}$, so that y^{-1} is contained in $\mathfrak{S}$, by Proposition 1. Thus we have proved that $\mathfrak{S}$ is a valuation ring. If $\mathfrak{R}$ is discrete, then, by the assumption $\dim_K (x) = \dim_{\mathrm{K}} (\xi)$, $\mathfrak{S}$ is also discrete.

Let K' be any extension of K and let $\mathfrak{R}'$ be an S-domain with the quotient field K' and the maximal ideal $\mathfrak{m}'$, such that $\mathfrak{m}' \cap \mathfrak{R} = \mathfrak{m}$, $\mathfrak{R}' \supset \mathfrak{R}$. When we fix a homomorphism of $\mathfrak{R}'$ into $\mathfrak{K}$ which is an extension of the natural mapping of $\mathfrak{R}$ onto K, we shall say that $\mathfrak{R}'$ is a *prolongation of* $\mathfrak{R}$ *in* K', and call the image of $\mathfrak{R}'$ by that homomorphism the *residue field* of $\mathfrak{R}'$. Thus, when we speak of a prolongation of an S-domain or a valuation ring, it should be remembered that "prolongation" is a combined conception of ring and homomorphism. Also we regard the specialization ring $[(x) \xrightarrow{\mathfrak{R}} (\xi)]$ with its natural mapping $(x) \longrightarrow (\xi)$ as a prolongation of $\mathfrak{R}$.

PROPOSITION 6. *Let* $\mathfrak{R}$ *be a discrete valuation ring and* $\mathfrak{R}'$ *a prolongation of* $\mathfrak{R}$ *in an extension* K' *of* K. *If* (ξ) *is a specialization of* (x) *over* $\mathfrak{R}$, *then there exists a generic specialization* (x') *of* (x) *over* K, *such that* (ξ) *is a specialization of* (x') *over* $\mathfrak{R}'$.

Proof. We may assume that (x) and (ξ) are both finite. Let $\mathfrak{A}$ be the ideal determined by (x) over the field K. We have a representation of the ideal $K'\mathfrak{A}$ of $K'[X]$ as an intersection of primary ideals $\mathfrak{Q}_i$, belonging respectively to prime ideals $\mathfrak{P}_i$. For every i, we have a set of quantities $(x^{(i)})$ such that $\mathfrak{P}_i$ is the ideal determined by $(x^{(i)})$ over K'. Then we have $(x) \xrightarrow{k} (x^{(i)})$. It is easy to prove that there exists, for every i, a generic specialization $(x'^{(i)})$ of (x) over K such that $(x^{(i)})$ is a specialization of $(x'^{(i)})$ over K'. Hence our proposition is proved if we show that for at least one value of i, $(x^{(i)})$ has (ξ) as a specialization over $\mathfrak{R}'$. Assume that this is impossible. Then there exists, for every i, a polynomial $F_i(X)$ in $\mathfrak{R}'[X]$ such that $F_i(x^{(i)}) = 0$ and that $\bar{F}_i(\xi) \neq 0$. Then $F_i(X) \in \mathfrak{P}_i$ for every i; so we can find an integer ρ such that the polynomial $F(X) = \prod_i F_i(X)^\rho$ is contained in $K'\mathfrak{A}$. As $F(X)$ is contained in $\mathfrak{R}'[X]$, we obtain an expression $F(X) = \sum_{\nu=1}^{s} b_\nu G_\nu(X)$ where $b_\nu \in \mathfrak{R}'$ and $G_\nu(X) \in \mathfrak{R}[X]$. Since $\mathfrak{R}$ is a discrete valuation ring, the finite $\mathfrak{R}$-module $\mathfrak{R}b_1 + \cdots + \mathfrak{R}b_s$ has an $\mathfrak{R}$-basis $(c_1, \cdots, c_r)$ such that $c_1, \cdots, c_r$ are linearly independent over K. Expressing b_ν in a form $b_\nu = \sum_\lambda u_{\nu\lambda} c_\lambda$ with $u_{\nu\lambda} \in \mathfrak{R}$, we have $F(X) = \sum_\nu b_\nu G_\nu(X) = \sum_\lambda c_\lambda H_\lambda(X)$, where $H_\lambda(X) = \sum_\nu u_{\nu\lambda} G_\nu(X) \in \mathfrak{R}[X]$. Then it follows from $F(X) \in K'\mathfrak{A}$ that $H_\lambda(X) \in \mathfrak{A}$ for every λ. Then we have $\bar{F}(\xi) = \sum_\lambda \bar{c}_\lambda \bar{H}_\lambda(\xi) = 0$; this is a contradiction.

PROPOSITION 7. *Let* $\mathfrak{R}$ *be a discrete valuation ring and* $\mathfrak{R}'$ *a prolongation of* $\mathfrak{R}$ *in an extension* K' *of* K. *Let* (x) *be a set of quantities such that*

$K(x)$ and K' are linearly disjoint over K, and (ξ) a finite specialization of (x) over $\mathfrak{R}$. Then (ξ) is a specialization of (x) over $\mathfrak{R}'$; and we have $[(x) \xrightarrow{\mathfrak{R}} (\xi)] = K(x) \cap [(x) \xrightarrow{\mathfrak{R}'} (\xi)]$.

Proof. Let $F(X)$ be a polynomial in $\mathfrak{R}'[X]$ such that $F(x) = 0$. We have an expression $F(X) = \sum_{\lambda=1}^{r} c_\lambda H_\lambda(X)$ where $c_\lambda \in \mathfrak{R}'$, $H_\lambda(X) \in \mathfrak{R}[X]$ for every λ. By the same procedure as in the proof of Proposition 6, we may assume that $c_1, \cdots, c_r$ are linearly independent over K. Then, by our assumption that K' and $K(x)$ are linearly disjoint over K, the relation $\sum_\lambda c_\lambda H_\lambda(x) = F(x) = 0$ implies that $H_\lambda(x) = 0$ for every λ. Since (ξ) is a specialization of (x) over $\mathfrak{R}$, we have $\bar{F}(\xi) = \sum_\lambda \bar{c}_\lambda \bar{H}_\lambda(\xi) = 0$; this proves the first assertion. Now let y be a quantity in $K(x) \cap [(x) \xrightarrow{\mathfrak{R}'} (\xi)]$; then we have an expression $y = P(x)/Q(x)$ with $P(X)$ and $Q(X)$ in $\mathfrak{R}'[X]$ and $\bar{Q}(\xi) \neq 0$. Let $(b_1, \cdots, b_s)$ be the set of all coefficients of P and Q; as in the proof of Proposition 6, we have an $\mathfrak{R}$-basis $d_1, \cdots, d_t$ of $\mathfrak{R}b_1 + \cdots + \mathfrak{R}b_s$ which are linearly independent over K. Then we have expressions $P(X) = \sum_\nu d_\nu F_\nu(X)$, $Q(X) = \sum_\nu d_\nu G_\nu(X)$ where the $F_\nu(X)$ and the $G_\nu(X)$ are in $\mathfrak{R}[X]$. By our assumption, from the relation $\sum_\nu d_\nu(F_\nu(x) - yG_\nu(x)) = 0$ follows that $F_\nu(x) = yG_\nu(x)$ for every ν. Since $\bar{Q}(\xi) \neq 0$, we have $\bar{G}_\nu(\xi) \neq 0$ for some ν; for that ν we have $y = F_\nu(x)/G_\nu(x)$; this proves the second assertion.

PROPOSITION 8. *Let $K(x)$ and $K(y)$ be two extensions of K which are linearly disjoint over K. If $\mathfrak{R}$ is a discrete valuation ring, then (ξ, η) is a specialization of (x, y) over $\mathfrak{R}$, if and only if (ξ) is a specialization of (x) over $\mathfrak{R}$ and (η) is a specialization of (y) over $\mathfrak{R}$.*

Proof. We shall prove only the direct part since the converse is obvious. We may assume that (ξ) and (η) are both finite. Put $K' = K(y)$ and $\mathfrak{R}' = [(y) \xrightarrow{\mathfrak{R}} (\eta)]$. We may consider $\mathfrak{R}'$ is a prolongation of $\mathfrak{R}$ in K'. Then our proposition is an immediate consequence of Proposition 7.

PROPOSITION 9. *Let $\mathfrak{R}$ be a discrete valuation ring. For any finitely generated extension K' of K, there exist a finitely generated extension K' of K and a prolongation $\mathfrak{R}'$ of $\mathfrak{R}$ in K' which is a discrete valuation ring, such that K' is the residue field of $\mathfrak{R}'$.*

Proof. Our proposition follows from Proposition 3 if K' is purely transcendental over K; so that we have only to prove our proposition in case where K' is a simple algebraic extension $\mathrm{K}(\xi)$ of K. Let $\bar{F}(X)=0$ be an irreducible equation for ξ over K; where we may assume that $\bar{F}$ is the class of a polynomial F in $\mathfrak{R}[X]$ of the same degree as $\bar{F}$, modulo $\mathfrak{m}$. Let x be a root of the equation $F(X)=0$ in $\bar{K}$, and put $K'=K(x)$. We have a discrete valuation ring $\mathfrak{R}'$ in K' with the maximal ideal $\mathfrak{m}'$, such that $\mathfrak{R}'\supset\mathfrak{R}$, $\mathfrak{m}'\cap\mathfrak{R}=\mathfrak{m}$. Since the class of x modulo $\mathfrak{m}'$ satisfies the equation $\bar{F}(X)=0$, we have $[\mathrm{K}':\mathrm{K}]\leqq[\mathfrak{R}'/\mathfrak{m}':\mathfrak{R}/\mathfrak{m}]$. It is well-known that $[\mathfrak{R}'/\mathfrak{m}':\mathfrak{R}/\mathfrak{m}]\leqq[K':K]$; so, by our definition of K', we have $[\mathrm{K}':\mathrm{K}]=[\mathfrak{R}'/\mathfrak{m}':\mathfrak{R}/\mathfrak{m}]=[K':K]$. This shows that $\mathfrak{R}'/\mathfrak{m}'$ is isomorphic to K' over K; so our proposition is proved.

2. Multiplicity of a proper specialization. As before, let $\mathfrak{R}$ be an S-domain with the quotient field K and the residue field K. We say that (η) is an *isolated specialization of* (y) *over* $\mathfrak{R}$, if whenever we have $(y)\xrightarrow{\mathfrak{R}}(\eta')$ and $(\eta')\xrightarrow{\mathrm{K}}(\eta)$, we also have $(\eta)\xrightarrow{\mathrm{K}}(\eta')$. A finite specialization of (y) over $\mathfrak{R}$ is called a *proper specialization of* (y) *over* $\mathfrak{R}$ if it is isolated and algebraic over K.

Let (y) be a set of algebraic quantities over K and let $(y^{(1)},\cdots,y^{(n)})$ be a complete set of conjugates of (y) over K. Let (η) be a specialization of (y) over $\mathfrak{R}$. If there exists a number μ, such that for every specialization $(\eta^{(1)},\cdots,\eta^{(n)})$ of $(y^{(1)},\cdots,y^{(n)})$ over $\mathfrak{R}$, the set (η) occurs exactly μ times among $(\eta^{(\nu)})$, then we call the number μ the *multiplicity of* (η) *as a specialization of* (y) *over* $\mathfrak{R}$. The main object of this section is to show that a proper specialization over a certain local domain has a well-defined multiplicity.

PROPOSITION 10. *Let $\mathfrak{R}$ be an integrally closed S-domain with the quotient field K, and (y) a set of algebraic quantities over K. If (y) is finite over $\mathfrak{R}$, then every specialization of (y) over $\mathfrak{R}$ is proper and has a multiplicity defined above.*

Proof. By Proposition 1 of Section 1, (y) is integral over $\mathfrak{R}$; so that every specialization of (y) over $\mathfrak{R}$ is finite algebraic over K. From this follows that every specialization of (y) over $\mathfrak{R}$ is proper. As for the assertion concerned with the multiplicity, first we shall prove it in case where the set (y) consists of a single quantity y. Let $F(Y)=0$ be the irreducible equation for y over K with the leading coefficient 1. Then $F(Y)$ is a polynomial contained in $\mathfrak{R}[Y]$ since $\mathfrak{R}$ is integrally closed. It can be easily verified that every specialization of the complete set of conjugates of y coincides with the set of all roots of the equation $\bar{F}(Y)=0$. This proves our proposition in the case when (y) consists of one quantity. To prove the general case we put $z = \sum_i t_i y_i$ where (t) is a set of independent variables over K. Let (τ) be a set of independent variables over K, and denote by $\mathfrak{S}$ the specialization ring $[(t) \xrightarrow{\mathfrak{R}} (\tau)]$. Then, by Proposition 4, $\mathfrak{S}$ is also an integrally closed S-domain; and it is easy to see that z is finite over $\mathfrak{S}$. If we denote by $(y^{(1)}, \cdots, y^{(n)})$ the complete set of conjugates of (y) over K and put $z_\nu = \sum_i t_i y_i^{(\nu)}$, then the set $(z_1, \cdots, z_n)$ consists of a repetition in a certain number of times of the complete set of conjugates of z over $K(t)$. So the general case is reduced to the case of one quantity.

Let $\mathfrak{R}$ be a local domain such that the completion $\mathfrak{R}^*$ of $\mathfrak{R}$ has no zero-divisor. We identify the quotient field K of $\mathfrak{R}$ with a subfield of the quotient field Ω of $\mathfrak{R}^*$ in the usual manner. The following two propositions are generalizations of Theorem 1 and Theorem 3 of [WF] Chapter III, Section 3. Proofs by the same method as in [WF] have been stated in Northcott [5] in a general form.

PROPOSITION 11. *Let $\mathfrak{R}^*$ be a complete local domain with the quotient field Ω and (y) a set of elements in some field containing Ω. If (y) has a proper specialization over $\mathfrak{R}^*$, then (y) is integral over $\mathfrak{R}^*$; and any two specializations of (y) over $\mathfrak{R}^*$ are conjugates of each other over the residue field of $\mathfrak{R}^*$.*

PROPOSITION 12. *Let $\mathfrak{R}$ be a local domain such that the completion $\mathfrak{R}^*$ of $\mathfrak{R}$ has no zero-divisor, (y) a set of generalized quantities which is algebraic over K, and (η) a specialization of (y) over $\mathfrak{R}$. Let Ω be the quotient field of $\mathfrak{R}^*$ and $\bar{\Omega}$ the algebraic closure of Ω. Then there exists an isomorphism of $K(y)$ over K into $\bar{\Omega}$ such that if (y') is the image of (y) we have $(y') \xrightarrow{\mathfrak{R}^*} (\eta)$.*

Now we consider specializations over a special type of local domains. Let $\mathfrak{o}$ be a discrete valuation ring with the quotient field k and the maximal ideal $\mathfrak{p}$, and let $\kappa = \mathfrak{o}/\mathfrak{p}$ be its residue field. Let $(x_1, \cdots, x_n)$ be a set of independent variables over k, and denote by $\mathfrak{R}$ the specialization ring $[(x_1, \cdots, x_n) \xrightarrow{\mathfrak{o}} (0, \cdots, 0)]$ and by $\mathfrak{R}^*$ its completion. Then $\mathfrak{R}^*$ is the ring of the formal power-series $\sum a_{(i)}^* x_1^{i_1} \cdots x_n^{i_n}$ with $a_{(i)}^*$ in the completion $\mathfrak{o}^*$ of $\mathfrak{o}$. The complete local domain $\mathfrak{R}^*$ is a regular local ring in Krull's sense[4] and is integrally closed.

PROPOSITION 13. *Let $\mathfrak{R}^*$ be the completion of the specialization ring $[(x_1, \cdots, x_n) \xrightarrow{\mathfrak{o}} (0, \cdots, 0)]$ in the above notations; let Ω be the quotient*

---

[4] Krull [3].

field of $\mathfrak{R}^*$ *and* (y) *a set of algebraic elements over* Ω. *If* (η) *is a proper specialization of* (y) *over* $\mathfrak{R}^*$, *then* (η) *has a multiplicity which is a multiple of* $[\kappa(\eta):\kappa]_i$.

Proof. By Proposition 11, (y) is integral over $\mathfrak{R}^*$; hence, by Proposition 10, every specialization (η') of (y) over $\mathfrak{R}^*$ has a multiplicity. Let $(y^{(1)},\cdots,y^{(d)})$ be a complete set of conjugates of (y) over Ω and $(\eta^{(1)},\cdots,\eta^{(d)})$ a specialization of $(y^{(1)},\cdots,y^{(d)})$ over $\mathfrak{R}^*$. By Proposition 11, the set $(\eta^{(1)},\cdots,\eta^{(d)})$ consists of conjugates of (η) over κ. Everyone of conjugates of (η) over κ is a specialization of (y) over $\mathfrak{R}^*$; so that it has a positive multiplicity as a specialization of (y) over $\mathfrak{R}^*$. Thus $(\eta^{(1)},\cdots,\eta^{(d)})$ contains all the conjugates of (η) over κ; and it is easy to see that all the conjugates of (η) occur in $(\eta^{(\nu)})$ in the same multiplicity. Hence our proposition is proved when we show that $d=[\Omega(y):\Omega]$ is a multiple of $[\kappa(\eta):\kappa]$. Let (t) be a set of n independent variables over k and (τ) a set of n independent variables over κ. We denote by $\mathfrak{R}_1$ the specialization ring $[(t)\xrightarrow{\mathfrak{o}}(\tau)]$ and by $\mathfrak{R}_1^*$ its completion. By Proposition 3 of Section 1, $\mathfrak{R}_1$ is a discrete valuation ring. We may regard the completion $\mathfrak{o}^*$ of $\mathfrak{o}$ as a subring of $\mathfrak{R}_1^*$. Now we shall define an isomorphic mapping f of $\mathfrak{R}^*$ over $\mathfrak{o}^*$ into $\mathfrak{R}_1^*$. For an element

$$A=\sum a_{(i)}{}^*x_1{}^{i_1}\cdots x_n{}^{i_n}\in\mathfrak{R}^*,$$

we define

$$f(A)=\sum a_{(i)}{}^*\pi^{i_1+\cdots+i_n}t_1{}^{i_1}\cdots t_n{}^{i_n}$$

where π is a prime element in $\mathfrak{o}$. It is easy to see that this mapping f is actually an isomorphic mapping into $\mathfrak{R}_1^*$. Then f is extended to an isomorphism of Ω into the quotient field Ω_1 of $\mathfrak{R}_1^*$, and is also extended to an isomorphism g of $\bar{\Omega}$ into $\bar{\Omega}_1$. We denote by (y') the image of (y) by g and by $(y'^{(1)},\cdots,y'^{(d)})$ the image of $(y^{(1)},\cdots,y^{(d)})$. If (ζ) is a specialization of $(y'^{(\nu)})$ over $\mathfrak{R}_1^*$, then (ζ) is a specialization of (y) over $\mathfrak{R}^*$; hence, by Proposition 11, (ζ) is a conjugate of (η) over κ. Now it is well-known that $[\Omega_1(y'^{(\nu)}):\Omega_1]$ is a multiple of $[\kappa(\tau,\zeta):\kappa(\tau)]$ since $\mathfrak{R}_1^*$ is a complete discrete valuation ring; the latter number is clearly equal to $[\kappa(\eta):\kappa]$, while $[\Omega(y):\Omega]$ is a sum of several $[\Omega_1(y'^{(\nu)}):\Omega_1]$; this completes our proof.

The following theorem is an extension of Theorem 4 of [WF] Chapter III, Section 4, and is basic for our whole theory.

THEOREM 2. *Let* (x) *be a set of* n *independent variables over* k *and* (y) *a set of algebraic quantities over* $k(x)$. *Let* (ξ) *be a finite specialization of* (x) *over* $\mathfrak{o}$ *and denote by* $\mathfrak{R}$ *the specialization ring* $[(x)\xrightarrow{\mathfrak{o}}(\xi)]$. *If* (η)

is a proper specialization of (y) *over* $\mathfrak{R}$, *then it has a multiplicity which is a multiple of* $[\kappa(\xi,\eta):\kappa(\xi)]_i$.[5]

Proof. We first assume that (ξ) is contained in κ. Then we have a set of quantities (a) in $\mathfrak{o}$ such that $(a)\xrightarrow{\mathfrak{o}}(\xi)$. Replacing the set (x_j) by (x_j-a_j), we may assume that $(\xi)=(0)$. In that special case, we can obtain our theorem by Proposition 12 and Proposition 13, in the same procedure as in the proof of Proposition 7 of [WF] Chapter III, Section 3. Now we shall prove the general case. Let s be the dimension of (ξ) over κ; if s is not 0, we may assume that $\xi_1,\cdots,\xi_s$ are independent variables over κ, and that (ξ) is algebraic over $\kappa(\xi_1,\cdots,\xi_s)$. By Proposition 3 of Section 1, the specialization ring $[(x_1,\cdots,x_s)\xrightarrow{\mathfrak{o}}(\xi_1,\cdots,\xi_s)]$ is a discrete valuation ring; if we denote it by $\mathfrak{o}_1$, we have $\mathfrak{R}=[(x_{s+1},\cdots,x_n)\xrightarrow{\mathfrak{o}_1}(\xi_{s+1},\cdots,\xi_n)]$. Hence, it is sufficient to prove our theorem in case where (ξ) is algebraic over κ. Suppose that (ξ) is algebraic over κ. By Proposition 9 of Section 1, we can find an algebraic extension k' of k and a valuation ring $\mathfrak{o}'$ which is a prolongation of $\mathfrak{o}$ in k', such that the residue field κ' of $\mathfrak{o}'$ is $\kappa(\xi)$. Let $(y^{(1)},\cdots,y^{(d)})$ be the complete set of conjugates of (y) over $k(x)$ and $(\eta^{(1)},\cdots,\eta^{(d)})$ a specialization of $(y^{(1)},\cdots,y^{(d)})$ over $\mathfrak{R}$. Then by Proposition 6 of Section 1, there exists a generic specialization $(x',y'^{(1)},\cdots,y'^{(d)})$ of $(x,y^{(1)},\cdots,y^{(d)})$ such that $(x',y'^{(1)},\cdots,y'^{(d)})\xrightarrow{\mathfrak{o}'}(\xi,\eta^{(1)},\cdots,\eta^{(d)})$. Since k' is algebraic over k, (x) is a generic specialization of (x') over k', and is extended to a generic specialization $(x,y''^{(1)},\cdots,y''^{(d)})$ of $(x',y'^{(1)},\cdots,y'^{(d)})$ over k'. Then $(x,y''^{(1)},\cdots,y''^{(d)})$ is a generic specialization of $(x,y^{(1)},\cdots,y^{(d)})$ over k and has the specialization $(\xi,\eta^{(1)},\cdots,\eta^{(d)})$ over $\mathfrak{o}'$. The set $(y''^{(1)},\cdots,y''^{(d)})$ is a permutation of $(y^{(1)},\cdots,y^{(d)})$ since $(y^{(1)},\cdots,y^{(d)})$ is the complete set of conjugates of (y) over $k(x)$. Therefore, it is sufficient to prove that for every specialization $(\eta^{(1)},\cdots,\eta^{(d)})$ of $(y^{(1)},\cdots,y^{(d)})$ over $(x)\xrightarrow{\mathfrak{o}'}(\xi)$ the set (η) occurs in $(\eta^{(1)},\cdots,\eta^{(d)})$ in the same times μ, and that μ is a multiple of $[\kappa(\xi,\eta):\kappa(\xi)]_i$. Now the set $(y^{(1)},\cdots,y^{(d)})$ consists of several complete sets of conjugates of sets $(y^{(\nu)})$ over $k'(x)$. Since $\kappa'=\kappa(\xi)$, we have $[\kappa(\xi,\eta):\kappa(\xi)]_i=[\kappa'(\eta):\kappa']_i$. Thus what we have to prove is reduced to the case in which our theorem is already proved.

The following three theorems are generalizations of Proposition 10, Proposition 11 and Theorem 5 in [WF] Chapter III, Section 4, and can be proved by the same arguments as there.

---

[5] This theorem holds for any regular local ring $\mathfrak{R}$. It can be easily verified that the specialization ring $[(x)\xrightarrow{\mathfrak{o}}(\xi)]$ in our theorem is a regular local ring.

THEOREM 3. *Let* (x) *be a set of independent variables over* k, (y) *and* (z) *two sets of algebraic quantities over* $k(x)$, *and* (ξ, η) *a finite specialization of* (x, y) *over* $\mathfrak{o}$. *Assume that* (η) *is a proper specialization, of multiplicity* μ, *of* (y) *over* $(x) \xrightarrow{\mathfrak{o}} (\xi)$, *and that* (z) *is finite over* $(x, y) \xrightarrow{\mathfrak{o}} (\xi, \eta)$. *Then the specializations* $(\zeta^{(\sigma)})$ *of* (z) *over* $(x, y) \xrightarrow{\mathfrak{o}} (\xi, \eta)$ *are in a finite number; for every* σ, $(\eta, \zeta^{(\sigma)})$ *is a proper specialization of* (y, z) *over* $(x) \xrightarrow{\mathfrak{o}} (\xi)$; *and if* μ_σ *is the multiplicity of that specialization, we have the relation*

$$\sum_{\sigma} \mu_\sigma = \mu[k(x, y, z) : k(x, y)].$$

THEOREM 4. *Let* (x) *and* (z) *be two sets of independent variables over* k; *let* (y) *be a set of algebraic quantities over* $k(x)$ *and* (w) *a set of algebraic quantities over* $k(z)$. *Assume that* $k(x, y)$ *and* $k(z, w)$ *are linearly disjoint over* k. *Let* (ξ, ζ) *be a finite specialization of* (x, z) *over* $\mathfrak{o}$; *let* (η) *be a proper specialization of* (y), *of multiplicity* μ, *over* $(x) \xrightarrow{\mathfrak{o}} (\xi)$, *and let* (ω) *be a proper specialization of* (w), *of multiplicity* ν, *over* $(z) \xrightarrow{\mathfrak{o}} (\zeta)$. *Then* (η, ω) *is a proper specialization of* (y, w) *over* $(x, z) \xrightarrow{\mathfrak{o}} (\xi, \zeta)$; *and the multiplicity of this specialization is equal to* $\mu\nu$.

THEOREM 5. *Let* (x) *be a set of independent variables over* k *and* (y) *a set of* m *algebraic quantities over* $k(x)$. *Let* (ξ, η) *be a finite specialization of* (x, y) *over* $\mathfrak{o}$, *and assume that there are* m *polynomials* $F_j(X, Y)$ *in* $\mathfrak{o}[X, Y]$, *such that* $F_j(x, y) = 0$ *for* $1 \leqq j \leqq m$ *and that the determinant* $|\partial F_j/\partial Y_h(\xi, \eta)|$ *is not* 0. *Then* (η) *is a proper specialization of* (y) *over* $(x) \xrightarrow{\mathfrak{o}} (\xi)$; *and its multiplicity is* 1.

PROPOSITION 14. *Let* (x) *be a set of independent variables over* k *and* (y) *a set of algebraic quantities over* $k(x)$. *Let* (ξ) *be a finite specialization of* (x) *over* $\mathfrak{o}$ *and* (η) *a proper specialization of* (y) *over* $(x) \xrightarrow{\mathfrak{o}} (\xi)$. *Then the following two conclusions hold.*

(I) *Let* (x') *be a finite specialization of* (x) *over* k *such that* (ξ) *is a specialization of* (x') *over* $\mathfrak{o}$. *Then there exists a specialization* (x', y') *of* (x, y) *over* k, *which has* (ξ, η) *as a specialization over* $\mathfrak{o}$.

(II) *Let* (ξ') *be a finite specialization of* (x) *over* $\mathfrak{o}$ *such that* (ξ) *is a specialization of* (ξ') *over* κ. *Then there exists a specialization* (ξ', η') *of* (x, y) *over* $\mathfrak{o}$, *which has* (ξ, η) *as a specialization over* κ.

This is also proved as in Proposition 12 of [WF] Chapter III, Section 4.

PROPOSITION 15. *Let* (x) *be a set of* n *quantities of dimension* r *over* k. *Let the* t_{ij} $(1 \leqq i \leqq r, 1 \leqq j \leqq n)$ *be* rn *independent variables over* $k(x)$

and the τ_{ij} $(1\leqq i\leqq r, 1\leqq j\leqq n)$ *rn independent variables over* κ. *Put* $\mathfrak{o}'=[(t)\xrightarrow{\mathfrak{o}}(\tau)]$, *and* $y_i=\sum_j t_{ij}x_j$ *for* $1\leqq i\leqq r$. *If* (ξ,η) *is a finite specialization of* (x,y) *over* $\mathfrak{o}'$, *then* (ξ) *is a proper specialization of* (x) *over* $(y)\xrightarrow{\mathfrak{o}'}(\eta)$.

Proof. It is enough to prove that if (ξ,η) is a finite specialization of (x,y) over $\mathfrak{o}'$ then (ξ) is algebraic over $\kappa(\tau,\eta)$. Let (ξ,η) be a finite specialization of (x,y) over $\mathfrak{o}'$ and s the dimension of (ξ) over κ. Then by Proposition 2 of Section 1, we have $s\leqq r$. Let (ξ') be a generic specialization of (ξ) over κ such that (ξ') and (τ) are independent over κ. Put $\eta_i'=\sum_j\tau_{ij}\xi_j'$ for $1\leqq i\leqq r$. Since $(\xi',\tau)\xrightarrow{\kappa}(\xi,\tau)$, we have $(\xi',\tau,\eta')\xrightarrow{\kappa}(\xi,\tau,\eta)$. Now by Proposition 24 of [WF] Chapter II, Section 5, (ξ') is finite over every finite specialization of $(\eta_1',\cdots,\eta_s')$ with reference to $\kappa(\tau_{11},\cdots,\tau_{sn})$. Hence (ξ) is algebraic over $\kappa(\tau,\eta)$.

PROPOSITION 16. *Let* (x), (y), (t), (τ) *and* $\mathfrak{o}'$ *be the same as in Proposition* 15. *Let the* s_j $(1\leqq j\leqq n)$ *be* n *independent variables over* $k(t,x)$ *and the* σ_j $(1\leqq j\leqq n)$ n *independent variables over* $\kappa(\tau)$. *Put* $\mathfrak{o}''=[(s)\xrightarrow{\mathfrak{o}'}(\sigma)]$ *and* $z=\sum_j s_jx_j$. *Let* (η,ζ) *be a finite specialization of* (y,z) *over* $\mathfrak{o}''$. *Then* ζ *is a proper specialization of* z *over* $(y)\xrightarrow{\mathfrak{o}''}(\eta)$; *and* (x) *is finite over* $(y,z)\xrightarrow{\mathfrak{o}''}(\eta,\zeta)$. *Let* (ξ) *be a specialization of* (x) *over* $(y,z)\xrightarrow{\mathfrak{o}''}(\eta,\zeta)$ *such that* (ξ) *and* (τ,σ) *are independent over* κ. *If* $k(x)$ *is separably generated over* k, *then we have* $k(t,s,x)=k(t,s,y,z)$; *and the multiplicity of* (ξ) *as a specialization of* (x) *over* $(y)\xrightarrow{\mathfrak{o}'}(\eta)$ *is equal to the multiplicity of* ζ *as a specialization of* z *over* $(y)\xrightarrow{\mathfrak{o}''}(\eta)$.

Proof. By the same method as in the proof of Proposition 24 of [WF] Chapter II, Section 5, we can prove that (x) is finite over $(y,z)\xrightarrow{\mathfrak{o}''}(\eta,\zeta)$. Let ξ be a specialization of (x) over $(y,z)\xrightarrow{\mathfrak{o}''}(\eta,\zeta)$. Then, by Proposition 15, (ξ) is algebraic over $\kappa(\tau,\eta)$; so that $\zeta=\sum_j\sigma_j\xi_j$ is algebraic over $\kappa(\tau,\sigma,\eta)$. This shows that every finite specialization of z over $(y)\xrightarrow{\mathfrak{o}''}(\eta)$ is algebraic over $\kappa(\tau,\sigma,\eta)$. Hence ζ is a proper specialization of z over $(y)\xrightarrow{\mathfrak{o}''}(\eta)$. Suppose that $k(x)$ is separably generated over k. Then $k(t,x)$ is separably algebraic over $k(t,y)$. Denote by $(x^{(1)},\cdots,x^{(d)})$ the complete set of conjugates of (x) over $k(t,y)$, and put $z^{(\nu)}=\sum_j s_jx_j^{(\nu)}$ for $1\leqq\nu\leqq d$; then it is easy to see that $(z^{(1)},\cdots,z^{(d)})$ is the complete set of conjugates of z over $k(t,s,y)$. This shows $[k(t,s,y,z):k(t,s,y)]=d=[k(t,s,x):k(t,s,y)]$; so we have $k(t,s,x)=k(t,s,y,z)$. Let (ξ) be a specialization of (x)

over $(y,z)\xrightarrow{\mathfrak{o}''}(\eta,\zeta)$ such that (ξ) and (τ,σ) are independent over κ. Let $(\xi^{(1)},\cdots,\xi^{(d)})$ be a specialization of $(x^{(1)},\cdots,x^{(d)})$ over $(y)\xrightarrow{\mathfrak{o}'}(\eta)$. We extend this specialization to a specialization

$$(x^{(1)},\cdots,x^{(d)},y,z^{(1)},\cdots,z^{(d)})\xrightarrow{\mathfrak{o}''}(\xi^{(1)},\cdots,\xi^{(d)},\eta,\zeta^{(1)},\cdots,\zeta^{(d)}).$$

If $(\xi^{(\nu)})=(\xi)$, then we have $\zeta^{(\nu)}=\sum_j\sigma_j\xi_j^{(\nu)}=\sum_j\sigma_j\xi_j=\zeta$. Conversely, if $\zeta^{(\nu)}=\zeta$, then we have $(x,y,z,t,s)\xrightarrow{k}(x^{(\nu)},y,z^{(\nu)},t,s)\xrightarrow{\mathfrak{o}}(\xi^{(\nu)},\eta,\zeta,\tau,\sigma)$. Since (x) is finite over $(y,z)\xrightarrow{\mathfrak{o}''}(\eta,\zeta)$, $(\xi^{(\nu)})$ is finite; so we have $\zeta=\sum_j\sigma_j\xi_j^{(\nu)}$. By Proposition 15, $(\xi^{(\nu)})$ is algebraic over $\kappa(\tau,\eta)$; since $\eta_i=\sum_j\tau_{ij}\xi_j$, $(\xi^{(\nu)})$ is also algebraic over $\kappa(\tau,\xi)$. By our assumption, (σ) is a set of independent variables over $\kappa(\tau,\xi)$; so that from $\sum_j\sigma_j\xi_j^{(\nu)}=\sum_j\sigma_j\xi_j$ follows $\xi_j^{(\nu)}=\xi_j$ $(1\leqq j\leqq n)$. Thus we have proved that $\zeta^{(\nu)}=\zeta$ if and only if $(\xi^{(\nu)})=(\xi)$. This proves the final assertion of our proposition.

By Proposition 14 and by Proposition 15, the following theorem can be proved in the same procedure as in the proof of Proposition 13 in [WF] Chapter III, Section 4.

THEOREM 6. *Let (x) and (y) be two sets of quantities, and (ξ,η) a finite specialization of (x,y) over $\mathfrak{o}$. Then if (η) is an isolated specialization of (y) over $(x)\xrightarrow{\mathfrak{o}}(\xi)$, the dimension of (η) over $\kappa(\xi)$ is at least equal to that of (y) over $k(x)$.*

PROPOSITION 17. *Let (y) be a set of independent variables over k and z an algebraic quantity over $k(y)$. Let $F(Y,Z)=0$ be an irreducible equation for (y,z) over k and assume that F is a primitive polynomial in $\mathfrak{o}[Y,Z]$. Let ζ be a proper specialization of z, of multiplicity 1, over a finite specialization $(y)\xrightarrow{\mathfrak{o}}(\eta)$. Then we have $\partial\bar{F}/\partial Z(\eta,\zeta)\neq 0$; and $[(y,z)\xrightarrow{\mathfrak{o}}(\eta,\zeta)]$ is integrally closed.*

This proposition is a translation of Proposition 19 in [WF] Chapter V, Section 3 to our case and can be proved by the same argument as there.

3. Reduction of affine varieties. Let k be a field with a discrete valuation, and $\mathfrak{o}$ be its valuation ring. We denote by $\mathfrak{p}$ the maximal ideal of $\mathfrak{o}$ and by κ the residue field $\mathfrak{o}/\mathfrak{p}$. We fix this field k and the valuation ring $\mathfrak{o}$; in the following we shall use these notations k, $\mathfrak{o}$, $\mathfrak{p}$ and κ always in this sense; and whenever we speak of a prolongation $\mathfrak{o}'$ of $\mathfrak{o}$ we presuppose that $\mathfrak{o}'$ is also a discrete valuation ring.

Let S^n and $\mathfrak{S}^n$ be the affine n-spaces defined with respect to the field k and κ, respectively. First we define the reduction of a variety as a point set.

Let U be a bunch[6] in S^n which is normally algebraic over k. We call a bunch attached to a set $\{(\alpha) \mid (\alpha) \varepsilon \mathfrak{S}^n, (a) \xrightarrow{\mathfrak{o}} (\alpha) \text{ for some } (a) \varepsilon U\}$ the *bunch obtained from U by the reduction with respect to* $\mathfrak{p}$ (or $\mathfrak{o}$), or the *reduced bunch,* and denote it by $\bar{U}$. It may happen that $\bar{U}$ is an empty set even if U is not empty.

THEOREM 7. *Let U be a bunch in S^n which is normally algebraic over k, and let $\mathfrak{a}$ be the ideal $\{F(X) \mid F(X) \varepsilon \mathfrak{o}[X], F(a) = 0 \text{ for every } (a) \varepsilon U\}$ in $\mathfrak{o}[X]$. Then, we have $\bar{U} = \{(\alpha) \mid (\alpha) \varepsilon \mathfrak{S}^n, \bar{F}(\alpha) = 0 \text{ for every } F(X) \varepsilon \mathfrak{a}\}$; so that $\bar{U}$ is a bunch in $\mathfrak{S}^n$ which is normally algebraic over κ. Furthermore if k' is any extension of k and $\mathfrak{o}'$ is a prolongation of $\mathfrak{o}$ in k', then we have $\bar{U} = \{(\alpha) \mid (\alpha) \varepsilon \mathfrak{S}^n, (a) \xrightarrow{\mathfrak{o}'} (\alpha) \text{ for some } (a) \varepsilon U\}$.*

Proof. By our definition of $\bar{U}$, it is obvious that if $(\alpha) \varepsilon \bar{U}$ then $\bar{F}(\alpha) = 0$ for every $F(X) \varepsilon \mathfrak{a}$. Now we express U as the union of prime rational cycles U_i over k: $U = U_1 \cup \cdots \cup U_h$. Let $(x^{(i)})$ be a generic point of U_i over k for every i; then we have $\mathfrak{a} = \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_h$ where

$$\mathfrak{a}_i = \{F(X) \mid F(X) \varepsilon \mathfrak{o}[X], F(x^{(i)}) = 0\}$$

for every i. Suppose that, for every i, (α) is not a specialization of $(x^{(i)})$ over $\mathfrak{o}$. Then there exists, for every i, a polynomial $F_i(X)$ in $\mathfrak{a}_i$ such that $\bar{F}_i(\alpha) \neq 0$. If we put $F(X) = \prod_i F_i(X)$ we have $F(X) \varepsilon \mathfrak{a}$ and $\bar{F}(\alpha) \neq 0$. This proves our first assertion. The second assertion follows from our definition of $\bar{U}$ and Proposition 6 of Section 1.

By this theorem the reduced bunch $\bar{U}$ is invariant under the extension of the basic field k and the prolongation of $\mathfrak{o}$ in that extension. This property of reduction also holds for reduction of cycles which is to be defined later.

PROPOSITION 18. *Let U and V be bunches which are normally algebraic over k. Then we have*

i) *If U and V are in the same ambient space S^n, then we have $\overline{U \cup V} = \bar{U} \cup \bar{V}$, $\overline{U \cap V} \subset \bar{U} \cap \bar{V}$.*

ii) *If U is in S^n and V is in S^m, then we have $\overline{U \times V} = \bar{U} \times \bar{V}$.*

Proof. The assertion i) is an immediate consequence of our definition. Now we shall prove ii). Let k' be a finite algebraic extension of k such that all the components of U and V are defined over k'. We have representations $U=\bigcup_i U_i$ and $V=\bigcup_j V_j$ where U_i and V_j are varieties defined over k'. Let $\mathfrak{o}'$ be a prolongation of $\mathfrak{o}$ in k', we consider the reduction with respect to $\mathfrak{o}'$. Since $\overline{U\times V}=\bigcup_{i,j}\overline{U_i\times V_j}$, it is sufficient to prove that $\overline{U_i\times V_j}=\overline{U}_i\times\overline{V}_j$ for every i and j. This is proved by Proposition 8 of Section 1 and by our definition.

---

[6] From now on we shall identify a bunch with the point set attached to that bunch. Also we shall use the term "prime rational cycle" for the bunch whose components are components of that cycle.

PROPOSITION 19. *Let U be a bunch in S^n which is normally algebraic over k and assume that the components of U are all of dimension r. If $\bar{U}$ is not empty, then the components of $\bar{U}$ are all of dimension r.*

This is an immediate consequence of Theorem 6 of Section 2.

Let A be a prime rational S^n-cycle over k. Now we shall define a rational $\mathfrak{S}^n$-cycle $\rho(A)$ over κ which will be called the cycle obtained from A by the reduction with respect to $\mathfrak{p}$ (or $\mathfrak{o}$).

LEMMA 2. *Let A^r be a prime rational S^n-cycle over a field K. Let the $F_i(X)=\sum_j c_{ij}X_j$ $(1\leqq i\leqq r)$ be r linearly independent linear forms in $K[X]$, and (v) a set of r independent variables over K; and let L^{n-r} be the variety defined by the set of equations $F_i(X)-v_i=0$ $(1\leqq i\leqq r)$. Then, if there is a point P in $A\cap L$, this point is algebraic over $K(v)$; and is a generic point of A over K; moreover the intersection product $U\cdot L$ is a prime rational cycle over $K(v)$ having P as a generic point over $K(v)$.*

This lemma is a generalization of Theorem 3 of [WF] Chapter V, Section 1, and follows immediately from that theorem.

Let A^r be a prime rational S^n-cycle over k. Suppose that $\bar{A}$ is not empty, and let $\mathfrak{A}_{11}$ be a component of $\bar{A}$. By Proposition 19, $\mathfrak{A}_{11}$ has the same dimension r as A. We shall define a certain multiplicity which will be denoted by $\mu(A,\mathfrak{A}_{11})$. Let

$$L^{n-r}=(\sum_j t_{ij}X_j-t_i=0\quad(1\leqq i\leqq r))$$

be a generic linear variety[7] over k in S^n; and let

$$\mathfrak{L}^{n-r}=(\sum\tau_{ij}X_j-\tau_i=0\quad(1\leqq i\leqq r))$$

be a generic linear variety over κ, in $\mathfrak{S}^n$. Then we have a point (x) in $L\cap A$, which is a generic point of A over $k(t_{ij})$, and we have a point (ξ) in

---

[7] By a linear variety $(\sum c_{ij}X_j-c_i=0\ (1\leqq i\leqq r))$ we understand the linear variety defined by the equations $\sum c_{ij}X_j-c_i=0$ $(1\leqq i\leqq r)$. We call it a generic linear variety over K if (c_i,c_{ij}) is a set of independent variables over K.

$\mathfrak{L} \cap \mathfrak{A}_{11}$ which is a generic point of $\mathfrak{A}_{11}$ over $\bar{\kappa}(\tau_{ij})$. The point (x) is algebraic over $k(t_i, t_{ij})$, and (ξ) is algebraic over $\kappa(\tau_i, \tau_{ij})$. Since $\mathfrak{A}_{11}$ is a component of $\bar{A}$, (ξ) is an isolated specialization of (x) over $\mathfrak{o}$. Hence (ξ) is a proper specialization of (x) over $(t_i, t_{ij}) \xrightarrow{\mathfrak{o}} (\tau_i, \tau_{ij})$; so that by Theorem 2 of Section 2, it has a multiplicity μ defined in Section 2. It is easy to see that this multiplicity μ depends only upon the cycle A and the component $\mathfrak{A}_{11}$ of A, and not upon the choices of (t), (τ), (x), and (ξ). We shall denote it by $\mu(A, \mathfrak{A}_{11})$. By Theorem 2, $\mu(A, \mathfrak{A}_{11})$, is a multiple of $[\kappa(\tau_i, \tau_{ij}, \xi) : \kappa(\tau_i, \tau_{ij})]_i$. It can be easily verified that the latter number is equal to the order of inseparability of $\mathfrak{A}_{11}$ over κ.[8] Let $\mathfrak{A}_1$ be the prime rational $\mathfrak{S}^n$-cycle over κ with the component $\mathfrak{A}_{11}$. We denote by $\mu(A, \mathfrak{A}_1)$ the integer $\mu(A, \mathfrak{A}_{11})/[\mathfrak{A}_{11} : \kappa]_i$. It is easy to see that $\mu(A, \mathfrak{A}_1)$ is actually determined only by A and $\mathfrak{A}_1$. We define a $\mathfrak{S}^n$-cycle $\rho(A)$ by $\rho(A) = \sum_{\nu} \mu(A, \mathfrak{A}_\nu)\mathfrak{A}_\nu$ where the sum is taken over all the prime rational $\mathfrak{S}^n$-cycles $\mathfrak{A}_\nu$ over κ whose components are components of $\bar{A}$. If $\bar{A}$ is empty, we put $\rho(A) = 0$. It is obvious that

$$\rho(A) = \sum_{\nu, i} \mu(A, \mathfrak{A}_{\nu i})\mathfrak{A}_{\nu i}$$

where the sum is taken over all the components $\mathfrak{A}_{\nu i}$ of $\bar{A}$. Let X^r be a rational S^n-cycle over k; and let $X = \sum a_\alpha A_\alpha$ be a representation of X as a linear combination of prime rational S^n-cycles A_α over k. We put $\rho(X) = \sum a_\alpha \rho(A_\alpha)$ and call it the *cycle obtained from X by the reduction with respect to* $\mathfrak{p}$ (or $\mathfrak{o}$). Then, ρ is an additive mapping of the group of rational S^n-cycles over k into the group of rational $\mathfrak{S}^n$-cycles over κ of the same dimension. We shall define in Section 4 such a mapping ρ also for cycles on an abstract variety.

PROPOSITION 20. *Let k' be an extension of k and $\mathfrak{o}'$ a prolongation of $\mathfrak{o}$ in k'. Let X^r be a rational S^n-cycle over k. If we define a cycle $\rho'(X)$ by the same procedure as above, with respect to $\mathfrak{o}'$, then $\rho'(X)$ coincides with $\rho(X)$.*

Proof. By linearity it is sufficient to prove our proposition in case where X is a prime rational cycle A over k. Let $A = \sum a_\alpha A_\alpha$ be a representation of A as a linear combination of prime rational cycles A over k'. Let

$$L^{n-r} = (\sum t_{ij}X_j - t_i = 0 \quad (1 \leqq i \leqq r))$$

be a generic linear variety over k' and

$$\mathfrak{L}^{n-r} = (\sum \tau_{ij}X_j - \tau_i = 0 \quad (1 \leqq i \leqq r))$$

---

[8] We shall denote by $[V : K]_i$ the order of inseparability of the variety V over a field K. If A is a prime rational cycle over K with the component V then we denote it also by $[A : K]_i$.

be a generic linear variety over the residue field κ' of $\mathfrak{o}'$. Let (x) be a point in $L \cap A$ and (ξ) a point in $\mathfrak{L} \cap \mathfrak{A}_{11}$, where $\mathfrak{A}_{11}$ is a component of $\bar{A}$. By Lemma 2, and by the equation $L \cdot A = \sum a_\alpha L \cdot A_\alpha$, the complete set of conjugates of (x) over $k(t_i, t_{ij})$ consists of the complete set of conjugates of $(x^{(\alpha)})$ over $k'(t_i, t_{ij})$, each set being repeated a_α times, where $(x^{(\alpha)})$ is a point in $L \cap A_\alpha$. By counting the number of (ξ) in a specialization of the complete set of conjugates of (x) over $k(t_i, t_{ij})$, over $(t_i, t_{ij}) \xrightarrow{\mathfrak{o}'} (\tau_i, \tau_{ij})$, we have $\mu(A, \mathfrak{A}_{11}) = \sum a_\alpha \mu'(A_\alpha, \mathfrak{A}_{11})$ where μ' is the multiplicity defined with respect to $\mathfrak{o}'$. By Theorem 7, a variety $\mathfrak{A}^r$ in $\mathfrak{S}^n$ is a component of $\rho(A)$ if and only if it is a component of $\rho'(A)$; so our proposition is proved.

PROPOSITION 21. *Let A^r be a prime rational S^n-cycle over k, and $\mathfrak{A}^r$ a prime rational $\mathfrak{S}^n$-cycle over κ with a component in $\bar{A}$. Let*

$$L^{n-r} = (\sum d_{ij} X_j - e_i = 0 \quad (1 \leqq i \leqq r))$$

be a linear variety in S^n, such that all the elements d_{ij} which are not in $\mathfrak{o}$ form a set of independent variables over k and that (e) is a set of independent variables over $k(d)$. Let

$$\mathfrak{L}^{n-r} = (\sum \delta_{ij} X_j - \epsilon_i = 0 \quad (1 \leqq i \leqq r))$$

be a linear variety in $\mathfrak{S}^n$, such that $\kappa(\delta)$ is regular over κ and that (ϵ) is a set of independent variables over $\kappa(\delta)$. Suppose that (δ, ϵ) is a specialization of (d, e) over $\mathfrak{o}$ and that there are a point (x) in $L \cap A$ and a point (ξ) in $\mathfrak{L} \cap \mathfrak{A}$. Then (ξ) is a proper specialization of (x) over $(d, e) \xrightarrow{\mathfrak{o}} (\delta, \epsilon)$, of multiplicity $\mu(A, \mathfrak{A}) \cdot [\kappa(\delta, \epsilon, \xi) : \kappa(\delta, \epsilon)]_i$.

Proof. First we remark that A is prime rational over $k(d)$ and that $\mathfrak{A}$ is prime rational over $\kappa(\delta)$. By Lemma 2, (x) is a generic point of A over $k(d)$ and (ξ) is a generic point of $\mathfrak{A}$ over $\kappa(\delta)$. Now it is obvious that (ξ) is a proper specialization of (x) over $(d, e) \xrightarrow{\mathfrak{o}} (\delta, \epsilon)$, since $(x) \longrightarrow (\xi)$ is isolated over $\mathfrak{o}$. By our assumption for the set (d) and by Theorem 2 of Section 2, (ξ) has a multiplicity μ as a specialization of (x) over $(d, e) \xrightarrow{\mathfrak{o}} (\delta, \epsilon)$. Let

$$\mathrm{M}^{n-r} = (\sum t_{ij} X_j - t_i = 0 \quad (1 \leqq i \leqq r))$$

be a generic linear variety over $k(d, e)$ and

$$\mathfrak{M}^{n-r} = (\sum \tau_{ij} X_j - \tau_i = 0 \quad (1 \leqq i \leqq r))$$

a generic linear variety over $\kappa(\delta, \epsilon)$. Then we have a point (y) in $M \cap A$ and a point (η) in $\mathfrak{M} \cap \mathfrak{A}$. Obviously, (η) is also a proper specialization of

(y) over $(t_{ij}, t_i) \xrightarrow{\mathfrak{o}} (\tau_{ij}, \tau_i)$; so that it has a multiplicity μ_1. We shall show $\mu = \mu_1$. Let $(y^{(1)}, \cdots, y^{(m)})$ be the complete set of conjugates of (y) over $k(t)$ and $(x^{(1)}, \cdots, x^{(m)})$ a specialization of $(y^{(1)}, \cdots, y^{(m)})$ over $(t_{ij}, t_i) \xrightarrow{k} (d_{ij}, e_i)$. Then, by Lemma 2 and the definition of intersection-multiplicity,[8′] $(x^{(1)}, \cdots, x^{(m)})$ consists of the complete set of conjugates of (x) over $k(d, e)$ and pseudo-points which may arise or may not arise. Let $(\xi^{(1)}, \cdots, \xi^{(m)})$ be a specialization of $(x^{(1)}, \cdots, x^{(m)})$ over $(d, e) \xrightarrow{\mathfrak{o}} (\delta, \epsilon)$; then we see that (ξ) occurs in $(\xi^{(1)}, \cdots, \xi^{(m)})$ exactly μ times. Thus we have shown $\mu = \mu_1$. Let $(\eta^{(1)}, \cdots, \eta^{(m)})$ be a specialization of $(y^{(1)}, \cdots, y^{(m)})$ over $(t_{ij}, t_i) \xrightarrow{\mathfrak{o}} (\tau_{ij}, \tau_i)$ and $(\xi'^{(1)}, \cdots, \xi'^{(m)})$, a specialization of $(\eta^{(1)}, \cdots, \eta^{(m)})$ over $(\tau_{ij}, \tau_i) \xrightarrow{\kappa} (\delta_{ij}, \epsilon_i)$. By definition of μ_1, we have (ξ) in $(\xi'^{(1)}, \cdots, \xi'^{(m)})$ exactly μ_1 times. If $(\xi'^{(\nu)}) = (\xi)$, then we have

$$(x) \xrightarrow{k} (y^{(\nu)}) \xrightarrow{\mathfrak{o}} (\eta^{(\nu)}) \xrightarrow{\kappa} (\xi).$$

By the isolatedness of $(x) \xrightarrow{\mathfrak{o}} (\xi)$, we have $(\xi) \xrightarrow{\kappa} (\eta^{(\nu)})$, this shows that $(\eta^{(\nu)})$ is in $\mathfrak{A}$. Since $(y^{(\nu)}, t_{ij}, t_i) \xrightarrow{\mathfrak{o}} (\eta^{(\nu)}, \tau_{ij}, \tau_i)$ and since $(y^{(\nu)}) \varepsilon M$, we see that $(\eta^{(\nu)})$ is in $\mathfrak{M} \cap \mathfrak{A}$, so that $(\eta^{(\nu)})$ is a conjugate of (η) over $\kappa(\tau)$. Now by our definition of $\mu(A, \mathfrak{A})$, the complete set of conjugates of (η) occurs in $(\eta^{(1)}, \cdots, \eta^{(m)})$ exactly $\mu(A, \mathfrak{A})$ times. Then, by Lemma 2 and by the definition of intersection-multiplicity, we have $\mu(A, \mathfrak{A})[\kappa(\delta, \epsilon, \xi) : \kappa(\delta, \epsilon)]_i =$ the number of (ξ) which occurs in $(\xi'^{(1)}, \cdots, \xi'^{(m)})$. This completes our proof.

THEOREM 8. *Let A^r be a rational S^m-cycle over k and B^s a rational S^n-cycle over k. Then we have $\rho(A \times B) = \rho(A) \times \rho(B)$.*

Proof. By linearity and by Proposition 20, it is sufficient to prove our theorem in case where A and B are varieties defined over k. Furthermore, we may assume that all the components of $\bar{A}$ and $\bar{B}$ are defined over κ, because for every extension κ' of κ, there exists, by Proposition 9 of Section 1, a prolongation $\mathfrak{o}'$ of $\mathfrak{o}$ such that κ' is the residue field of $\mathfrak{o}'$. By Proposition 18, a variety $\mathfrak{C}$ in $\mathfrak{S}^n \times \mathfrak{S}^m$ is a component of $\overline{A \times B}$ if and only if it is a product $\mathfrak{A} \times \mathfrak{B}$ where $\mathfrak{A}$ is a component of $\bar{A}$ and $\mathfrak{B}$ is a component of $\bar{B}$. Hence it suffices to prove $\mu(A \times B, \mathfrak{A} \times \mathfrak{B}) = \mu(A, \mathfrak{A}) \cdot \mu(B, \mathfrak{B})$. Let $M^{m-r} = (\sum t_{ij}X_j - t_i = 0 \ \ (1 \leqq i \leqq r))$ be a generic linear variety over k in S^m and $N^{n-s} = (\sum z_{ij}X_j - z_i = 0 \ \ (1 \leqq i \leqq s))$ a generic linear variety over $k(t)$ in S^n. Correspondingly, let $\mathfrak{M}^{m-r} = (\sum \tau_{ij}X_j - \tau_i = 0 \ (1 \leqq i \leqq r))$ be a generic linear variety over κ in $\mathfrak{S}^m$ and $\mathfrak{N}^{n-s} = (\sum \zeta_{ij}X_j - \zeta_i = 0 \ (1 \leqq i \leqq s))$ a generic linear variety over $\kappa(\tau)$ in $\mathfrak{S}^n$. We have points (x),

---

[8′] See [WF] Chapter V, Section 1.

(y), (ξ) and (η) respectively in $M \cap A$, $N \cap B$, $\mathfrak{M} \cap \mathfrak{A}$ and $\mathfrak{N} \cap \mathfrak{B}$. Applying Proposition 21 to $A \times B$, $\mathfrak{A} \times \mathfrak{B}$, $M \times N$ and $\mathfrak{M} \times \mathfrak{N}$, we see that the multiplicity of (ξ, η) as a specialization of (x, y) over $(t_{ij}, z_{ij}, t_i, z_i) \xrightarrow{\mathfrak{o}} (\tau_{ij}, \zeta_{ij}, \tau_i, \zeta_i)$ is equal to $\mu(A \times B, \mathfrak{A} \times \mathfrak{B})$. By Theorem 4 of Section 2 and by our definition of $\mu(A, \mathfrak{A})$, this multiplicity is equal to $\mu(A, \mathfrak{A}) \cdot \mu(B, \mathfrak{B})$.

Now we consider the relation between our multiplicity $\mu(A, \mathfrak{A})$ and the intersection-multiplicity.

Let U^n be a variety defined over a field K; let A^r and B^s be prime rational U-cycles over K. If there is a component C_1 of the bunch $A \cap B$, which is simple on U and has dimension $r + s - n$, then every conjugate of C_1 over K is also a simple component of $A \cap B$ on U with the dimension $r + s - n$. Consider the number

$$[A:K]_i[B:K]_i[C_1:K]_i^{-1} \sum_{p,q} i(A_p \cdot B_q, C_1; U)$$

where the sum is taken over all the A_p and all the B_q which are respectively components of A and B, and which contains C_1. It is easy to see that this number is invariant when we replace C_1 with any conjugate of C_1 over K. We denote this number by $i_K(A \cdot B, C; U)$ where C is a prime rational U-cycle over K with the component C_1. We shall say that C is a *prime component of* $A \cap B$ on U, over K, if U, A, B and C are in the above situation. We shall also use the notation $j_K(V \cdot L, W)$ for an intersection of a prime rational cycle and a linear variety.

THEOREM 9. *Let V^r be a variety defined over k in S^n. Let*

$$L^{n-s} = (\sum c_{ij}X_j - c_i = 0 \quad (1 \leqq i \leqq s))$$

be a linear variety in S^n where $c_{ij} \in \mathfrak{o}$, $c_i \in \mathfrak{o}$; and let

$$\mathfrak{L}^{n-s} = (\sum \gamma_{ij}X_j - \gamma_i = 0 \quad (1 \leqq i \leqq s))$$

be a linear variety in $\mathfrak{S}^n$ such that $(c_{ij}, c_i) \xrightarrow{\mathfrak{o}} (\gamma_{ij}, \gamma_i)$. If there is a component $\mathfrak{W}_1$ of dimension $r - s$ in $\bar{V} \cap \bar{L}$, then there exists a prime rational cycle W^{r-s} over k with a component in $V \cap L$ such that $\mathfrak{W}_1$ is a component of $\bar{W}$. If we denote by the W_ν such cycles, then for every ν, W_ν is a prime component of $V \cap L$ over k; and we have

$$\sum_\nu j_k(V \cdot L, W_\nu)\mu(W_\nu, \mathfrak{W}) = \sum_\nu \mu(V, \mathfrak{V}_\nu) j_\kappa(\mathfrak{V}_\nu \cdot \mathfrak{L}, \mathfrak{W}),$$

where $\mathfrak{W}$ is the prime rational cycle over κ with the component $\mathfrak{W}_1$, and the $\mathfrak{V}_\nu$ are all the prime components of $\rho(V)$[9] *over κ which contain $\mathfrak{W}_1$.*

---

[9] We understand by a prime component over K, of a rational cycle over K, a prime rational cycle over K with a component which appears in the reduced expression of that rational cycle.

Proof. Let $M^{n-s} = (\sum t_{ij}X_j - t_i = 0 \ (1 \leqq i \leqq s))$ be a generic linear variety over k and $N^{n-r+s} = (\sum z_{ij}X_j - z_i = 0 \ (1 \leqq i \leqq r-s))$ a generic linear variety over $k(t)$. Correspondingly, let $\mathfrak{M}^{n-s} = (\sum \tau_{ij}X_j - \tau_i = 0$ $(1 \leqq i \leqq s))$ be a generic linear variety over κ and $\mathfrak{N}^{n-r+s} = (\sum \zeta_{ij}X_j - \zeta_i = 0$ $(1 \leqq i \leqq r-s))$ a generic linear variety over $\kappa(\tau)$. Then we have a point (x) in $V \cap M \cap N$ and a point (η) in $\mathfrak{W}_1 \cap \mathfrak{N}$. By Lemma 2, (x) is a generic point of V over $k(t_{ij}, z_{ij})$; and (η) is a generic point of $\mathfrak{W}_1$ over $\bar{\kappa}(\zeta_{ij})$. It is easy to see that (η) is a specialization of (x) over $(t_{ij}, t_i, z_{ij}, z_i) \xrightarrow{\mathfrak{o}} (\gamma_{ij}, \gamma_i, \zeta_{ij}, \zeta_i)$. Moreover (η) is an isolated specialization of (x) over $(t_{ij}, t_i) \xrightarrow{\mathfrak{o}} (\gamma_{ij}, \gamma_i)$. For, if $(x, t) \xrightarrow{\mathfrak{o}} (\eta', \gamma) \xrightarrow{\kappa} (\eta, \gamma)$, then (η') is in $\bar{V} \cap \bar{L}$ and has (η) as a specialization over κ. Since (η) is a generic point of a component $\mathfrak{W}_1$ of $\bar{V} \cap \bar{L}$, we have $(\eta) \xrightarrow{\kappa} (\eta')$. Hence (η) is a proper specialization of (x) over $(t, z) \xrightarrow{\mathfrak{o}} (\gamma, \zeta)$, and by Theorem 2 of Section 2, has a multiplicity μ. Now consider a series of specializations $(t, z) \xrightarrow{k} (c, z) \xrightarrow{\mathfrak{o}} (\gamma, \zeta)$. By (I) of Proposition 14 of Section 2, there exists a set (y) such that $(x, t, z) \xrightarrow{k} (y, c, z) \xrightarrow{\mathfrak{o}} (\eta, \gamma, \zeta)$. Since $\sum t_{ij}x_j = t_i \ (1 \leqq i \leqq s)$, we have $\sum c_{ij}y_j = c_i \ (1 \leqq i \leqq s)$; so that (y) is contained in $V \cap L$. Let W_0 be the locus of (y) over k, and W_0' the component of $V \cap L$ which contains W_0. Then we have $\bar{W}' \supset \mathfrak{W}_1$ where W' is the prime rational cycle over k, with the component W_0'. Since $\mathfrak{W}_1$ is a component of $\bar{V} \cap \bar{L}$, $\mathfrak{W}_1$ must be a component of $\bar{W}'$; so that we have $\dim \mathfrak{W}_1 = \dim W' \geqq \dim W_0 \geqq \dim \mathfrak{W}_1$. From this follows that W_0 is a component of $V \cap L$ and that $\dim W_0 = \dim \mathfrak{W}_1$. This proves the existence of a prime rational cycle W over k in our theorem. Let the W_ν be such cycles; then by the same argument as above we have $\dim W_\nu = \dim \mathfrak{W}_1 = r + s - n$; so that the W_ν are prime components of $V \cap L$ over k. Let k' be an extension of k and $\mathfrak{o}'$ a prolongation of $\mathfrak{o}$ in k' such that all the components of $V \cap L$ are defined over k' and that all the components of $\bar{V}$ and $\bar{V} \cap \bar{L}$ are defined over the residue field κ' of $\mathfrak{o}'$.[9'] We denote by $\mu'(A, \mathfrak{A})$ the multiplicity with respect to $\mathfrak{o}'$. Denote by the $W_{\nu\lambda}$ the components of W_ν for every ν and by the $\mathfrak{B}_{\nu\lambda}$ the components of $\mathfrak{B}_\nu$ for every ν. By Proposition 20, we have

$$[\mathfrak{W} : k]_i \sum_\nu j_k(V \cdot L, W_\nu)\mu(W_\nu, \mathfrak{W})$$
$$= \sum_\nu j(V \cdot L, W_{\nu 1})[W_\nu : k]_i^{-1} \sum_\lambda [W_\nu : k]_i \mu'(W_{\nu\lambda}, \mathfrak{W}_1)$$
$$= \sum_{\nu,\lambda} j(V \cdot L, W_{\nu\lambda})\mu'(W_{\nu\lambda}, \mathfrak{W}_1).$$

[9'] The existences of such an extension k' and a prolongation $\mathfrak{o}'$ are assured by Proposition 9 in Section 1.

Similarly we have

$$[\mathfrak{W}:k]_i \sum_{\nu} \mu(V,\mathfrak{B}_\nu) j_\kappa(\mathfrak{B}_\nu \cdot \mathfrak{L}, \mathfrak{W}) = \sum_{\nu,\lambda} \mu'(V,\mathfrak{B}_{\nu\lambda}) j(\mathfrak{B}_{\nu\lambda} \cdot \mathfrak{L}, \mathfrak{W}_1).$$

By these equations it suffices to prove the formula in our theorem in case where all the components of $V \cap L$ are defined over k and where all the components of $\bar{V}$ and $\bar{V} \cap \bar{L}$ are defined over κ. Assume that we are in this situation. Let (y') be a point in $W_\nu \cap N$, then (y') is a generic point of W_ν over k; and we have $(x,t,z) \xrightarrow{k} (y',c,z) \xrightarrow{\mathfrak{o}} (\eta,\gamma,\zeta)$. Conversely if we have $(x,t,z) \xrightarrow{k} (y',c,z) \xrightarrow{\mathfrak{o}} (\eta,\gamma,\zeta)$, then (y') is a generic point of some W_ν and is contained in $W_\nu \cap N$. This is already proved. Any such two points (y') and (y'') are conjugates of each other over $k(z)$ if and only if the loci of them are the same. Now let $(x^{(1)},\cdots,x^{(d)})$ be the complete set of conjugates of (x) over $k(t,z)$. We extend the specializations $(t,z) \xrightarrow{k} (c,z) \xrightarrow{\mathfrak{o}} (\gamma,\zeta)$ to specializations

$$(x^{(1)},\cdots,x^{(d)},t,z) \xrightarrow{k} (y^{(1)},\cdots,y^{(d)},c,z) \xrightarrow{\mathfrak{o}} (\eta^{(1)},\cdots,\eta^{(d)},\gamma,\zeta).$$

If $(\eta^{(\lambda)}) = (\eta)$, then $(y^{(\lambda)})$ is in $W_\nu \cap N$ for some W_ν. By the definition of $j(V \cdot L, W_\nu)$ ([WF] Chapter V, Section 2), we have the complete set of conjugates of $(y^{(\lambda)})$ over $k(z)$ in $(y^{(1)},\cdots,y^{(d)})$, exactly $j(V \cdot L, W_\nu)$ times. Hence we have $\sum_{\nu} j(V \cdot L, W_\nu)\mu(W_\nu,\mathfrak{W})$ as the number μ of (η) which occurs in $(\eta^{(1)},\cdots,\eta^{(d)})$. We consider another series of specializations $(t,z) \xrightarrow{\mathfrak{o}} (\tau,\zeta) \xrightarrow{\kappa} (\gamma,\zeta)$. Then a point (ξ) is a generic point of one of the $\mathfrak{B}_\nu$ and is contained in $\mathfrak{M} \cap \mathfrak{N}$ if and only if

$$(x,t,z) \xrightarrow{\mathfrak{o}} (\xi,\tau,\zeta) \xrightarrow{\kappa} (\eta,\gamma,\zeta).$$

The loci of any such two points (ξ') and (ξ'') are the same if and only if they are conjugates of each other over $\kappa(\tau,\zeta)$. Hence we have as above, $\mu = \sum_{\nu} \mu(V,\mathfrak{B}_\nu) j(\mathfrak{B}_\nu \cdot \mathfrak{L}, \mathfrak{W})$. This completes our proof.

We denote by $j_\kappa(V \cdot \mathfrak{L}, \mathfrak{W})$ the number $\sum_{\nu} \mu(V,\mathfrak{B}_\nu) j_\kappa(\mathfrak{B}_\nu \cdot \mathfrak{L}, \mathfrak{W})$ in the above theorem. It is easily verified that the number $j_\kappa(V \cdot \mathfrak{L}, \mathfrak{W})[\mathfrak{W}:\kappa]_i$ is equal to the multiplicity of (η) as a specialization of (x) over $(t,z) \xrightarrow{\mathfrak{o}} (\gamma,\zeta)$ where the notations are the same as in the above proof. We shall also denote by $j(V \cdot \mathfrak{L}, \mathfrak{W}_1)$ this number $j_\kappa(V \cdot \mathfrak{L}, \mathfrak{W})[\mathfrak{W}:\kappa]_i$ where $\mathfrak{W}_1$ is a component of $\mathfrak{W}$. Using this j, we have $\mu(V,\mathfrak{B}) = j(V \cdot \mathfrak{S}^n, \mathfrak{B})$ for every prime component $\mathfrak{B}$ of $\rho(V)$.

Before we deal with intersections in the general case, we extend the conception of the simple point. Let V^r be a variety defined over k in S^n, and denote by $\mathfrak{a}$ the ideal $\{F(X)\,|\,F(X)\,\varepsilon\,\mathfrak{o}[X], F(a)=0$ for every $(a)\,\varepsilon\,V\}$ in $\mathfrak{o}[X]$. We say that a point (α) in $\bar{V}$ is *simple on* V when there exists a set of polynomials $F_i(X)$ $(1 \leqq i \leqq n-r)$ in $\mathfrak{a}$ such that rank $\|\partial F_i/\partial X_j(\alpha)\|$ is equal to $n-r$. We call the linear variety $(\sum \partial F_i/\partial X_j(\alpha)(X_j-\alpha_j)=0$ $(1\leqq i \leqq n-r))$ in $\mathfrak{S}^n$ the *tangent linear variety to* V *at* (α). Also we say that a variety $\mathfrak{W}$ contained in $\bar{V}$ is *simple on* V if some point on $\mathfrak{W}$ is simple on V.

The following proposition is proved as in Theorem 5 of [WF] Chapter IV, Section 3.

PROPOSITION 22. *Let* V *and* W *be varieties defined over* k, *respectively in* S^n *and* S^m. *Let* (α) *be a point in* $\bar{V}$ *and* (β) *a point in* $\bar{W}$. *Then the point* $(\alpha)\times(\beta)$ *is simple on* $V\times W$ *if and only if* (α) *is simple on* V *and* (β) *is simple on* W.

Let $\mathfrak{A}$ be a variety in $\bar{V}$ and suppose that $\mathfrak{A}$ is simple on V. We shall use the term "*uniformizing set of linear forms for* V *along* $\mathfrak{A}$." We understand, by this term, a set of r linear forms $\Phi_i(X)=\sum \gamma_{ij}X_j$ $(1\leqq i\leqq r)$ such that the linear variety $(\Phi_i(X)=0$ $(1\leqq i\leqq r))$ in $\mathfrak{S}^n$ is of dimension $n-r$ and is transversal to the tangent linear variety to V at some point of $\mathfrak{A}$. This is an analogy of that for the ordinary case in [WF] Chapter IV, Section 6.

THEOREM 10. *Let* V^r *be a variety in* S^n *defined over* k *and* (α) *a point in* $\bar{V}$ *which is simple on* V. *Let* $\mathfrak{L}^{n-r}=(\sum\gamma_{ij}X_j-\gamma_i=0$ $(1\leqq i\leqq r))$ *be a linear variety in* $\mathfrak{S}^n$ *which is transversal to the tangent linear variety* $\mathfrak{T}^r$ *to* V *at* (α). *Suppose that* (α) *is rational over* κ *and that* (γ_{ij},γ_i) *is in* κ. *Then* (α) *is a component of* $\bar{V}\cap\mathfrak{L}$; *and we have* $j(V\cdot\mathfrak{L},(\alpha))=1$. *In particular, this shows that* (α) *is contained in only one component* $\mathfrak{B}$ *of* $\bar{V}$. *Moreover* (α) *is a simple point on* $\mathfrak{B}$; *and* $\mathfrak{T}$ *is the tangent linear variety to* $\mathfrak{B}$ *at* (α).

Proof. By Theorem 5 of Section 2, and by the same argument as in the proof of Proposition 7 of [WF] Chapter V, Section 1, it can be proved that $j(V\cdot\mathfrak{L},(\alpha))=1$. Then if we denote by $\mathfrak{B}_1,\cdots,\mathfrak{B}_h$ all the components of V which contain (α), we have

$$1=\sum_{\nu=1}^{h}\mu(V,\mathfrak{B}_\nu)\,j(\mathfrak{B}_\nu\cdot\mathfrak{L},(\alpha)).$$

Hence h must be equal to 1. The remaining part of our theorem is obvious.

COROLLARY. *Let* V *be a variety defined over* k *and* $\mathfrak{B}$ *a component of* by the above theorem and the proof of it. From this follows $[\mathfrak{B}:\kappa]_i=1$.

Proof. By our definition, there exists a point (α) on $\mathfrak{B}$ which is simple on V. By Proposition 9 of Section 1, there exist an extension k' of k and a prolongation $\mathfrak{o}'$ of $\mathfrak{o}$ in k' such that the residue field of $\mathfrak{o}'$ is $\kappa(\alpha)$. Since $\mu(V, \mathfrak{B})$ is invariant when considered with respect to $\mathfrak{o}'$, we have $\mu(V, \mathfrak{B}) = 1$ by the above theorem and the proof of it. From this follows $[\mathfrak{B} : \kappa]_i = 1$.

THEOREM 11. *Let U^n be a variety defined over k; let A^r and B^r be two prime rational U-cycles over k. Suppose that there exists a component $\mathfrak{C}_1^{r+s-n}$ of $\bar{A} \cap \bar{B}$ which is simple on U. Then, there exists a prime rational cycle C over k with a component in $A \cap B$ such that $\mathfrak{C}_1$ is a component of $\bar{C}$. If we denote by the C_ν such cycles, then every C_ν is a prime component of $A \cap B$ on U over k. Let the $\mathfrak{A}_p{}^r$ and the $\mathfrak{B}_q{}^s$ be those prime rational cycles over κ which contain $\mathfrak{C}_1$, and of which components are components of $\bar{A}$ and $\bar{B}$, respectively; and let $\mathfrak{U}$ be the component of U which contains $\mathfrak{C}_1$ (by Theorem 10, such $\mathfrak{U}$ is uniquely determined), and $\mathfrak{C}$ the prime rational cycle over κ with the component $\mathfrak{C}_1$. If $\mathfrak{U}$ is defined over κ, then, the $\mathfrak{A}_p$ and the $\mathfrak{B}_q$ are all contained in $\mathfrak{U}$; and $\mathfrak{C}$ is a prime component of $\mathfrak{A}_p \cap \mathfrak{B}_q$ on $\mathfrak{U}$ over κ, for every p and q; and we have the following formula:*

$$\sum_{\nu} i_k(A \cdot B, C_\nu; U)\mu(C_\nu, \mathfrak{C}) = \sum_{p,q} \mu(A, \mathfrak{A}_p)\mu(B, \mathfrak{B}_q) i_\kappa(\mathfrak{A}_p \cdot \mathfrak{B}_q, \mathfrak{C}; \mathfrak{U}).$$

Proof. Let k' be an extension of k and $\mathfrak{o}'$ a prolongation of $\mathfrak{o}$ in k' such that all the components of A, B and $A \cap B$ are defined over k', and that all the components of $\bar{A}$, B and $\bar{A} \cap B$ are defined over the residue field κ' of $\mathfrak{o}'$. We shall denote by $\mu'(A, \mathfrak{A})$ our multiplicity with respect to $\mathfrak{o}'$. Suppose that there exists a component C_0 of $A \cap B$ such that $\bar{C}_0 \supset \mathfrak{C}_1$, and denote by C the prime rational cycle over k with the component C_0. Then, similarly as in the proof of Theorem 9, we can prove that $\mathfrak{C}_1$ is a component of $\bar{C}$ and that C is a prime component of $A \cap B$ on U over k. If we define the C_ν, the $\mathfrak{A}_p$, the $\mathfrak{B}_q$ and $\mathfrak{U}$ as in our theorem, then it is obvious that the $\mathfrak{A}_p$ and the $\mathfrak{B}_q$ are all contained in $\mathfrak{U}$ and that $\mathfrak{C}$ is a prime component of $\mathfrak{A}_p \cap \mathfrak{B}_q$ on $\mathfrak{U}$ over κ. Denote the components of A, B, C_ν, $\mathfrak{A}_p$ and $\mathfrak{B}_q$ respectively by A_α, B_β, $C_{\nu\lambda}$, $\mathfrak{A}_{pu}$ and $\mathfrak{B}_{qv}$. By Proposition 20 and by our definition of $i_k(A \cdot B, C_\nu; U)$, we have

$$[A : k]_i^{-1}[B : k]_i^{-1}[\mathfrak{C} : \kappa]_i \sum_{\nu} i_k(A \cdot B, C_\nu; U)\mu(C_\nu, \mathfrak{C}) = \sum_{\alpha,\beta,\nu,\lambda} i(A_\alpha \cdot B_\beta, C_{\nu\lambda}; U)\mu'(C_{\nu\lambda}, \mathfrak{C}_1)$$

and

$$[A : k]_i^{-1}[B : k]_i^{-1}[\mathfrak{C} : \kappa]_i \sum_{p,q} \mu(A, \mathfrak{A}_p)\mu(B, \mathfrak{B}_q) i_\kappa(\mathfrak{A}_p \cdot \mathfrak{B}_q, \mathfrak{C}; \mathfrak{U}) = \sum_{\alpha,\beta,p,q,u,v} \mu'(A_\alpha, \mathfrak{A}_{pu})\mu'(B_\beta, \mathfrak{B}_{qv}) i(\mathfrak{A}_{pu} \cdot \mathfrak{B}_{qv}, \mathfrak{C}_1; \mathfrak{U}).$$

By these equations, it is sufficient to prove our theorem in case where A, B, and the components of $A \cap B$ are all defined over k and where all the components of $\bar{A}$, B and $\bar{A} \cap \bar{B}$ are defined over κ. Assume that we are in this case. Let S^N and $\mathfrak{S}^N$ be respectively the ambient spaces for U and $\mathfrak{U}$. Let $\bar{F}_i(X)$ $(1 \leqq i \leqq n)$ be a uniformizing set of linear forms for U along $\mathfrak{C}$. We may assume that for every i, $\bar{F}_i(X)$ is the class of a linear form $F_i(X)$ in $\mathfrak{o}[X]$ modulo $\mathfrak{p}$. We denote by Λ^{2N-n} and $\bar{\Lambda}^{2N-n}$ respectively the linear variety $(F_i(X-X') = 0 \ (1 \leqq i \leqq n))$ in $S^N \times S^N$ and the linear variety $(\bar{F}_i(X-X') = 0 \ (1 \leqq i \leqq n)$ in $\mathfrak{S}^N \times \mathfrak{S}^N$. Let (γ) be a generic point of $\mathfrak{C}$ over κ and $\Delta_{\mathfrak{C}}$ the locus of (γ, γ) over κ. Then, $\Delta_{\mathfrak{C}}$ is a proper component of $\bar{\Lambda} \cap (\mathfrak{A}_p \times \mathfrak{B}_q)$; and by definition of i ([WF] Chapter VI, Section 1), we have $i(\mathfrak{A}_p \cdot \mathfrak{B}_q, \mathfrak{C}; \mathfrak{U}) = j[(\mathfrak{A}_p \times \mathfrak{B}_q) \cdot \bar{\Lambda}, \Delta_{\mathfrak{C}}]$. By Theorem 9, there exists a component D of $\Lambda \cap (A \times B)$ such that $\bar{D} \supset \Delta_{\mathfrak{C}}$. (We may assume that the components of $\Lambda \cap (A \times B)$ are all defined over k.) If we denote by $D_1, \cdots, D_h$ such components of $\Lambda \cap (A \times B)$, then by the same theorem, we have

$$\sum_{\nu} j[(A \times B) \cdot \Lambda, D_\nu] \mu(D_\nu, \Delta_{\mathfrak{C}}) = \sum_{p,q} \mu(A \times B, \mathfrak{A}_p \times \mathfrak{B}_q) j[(\mathfrak{A}_p \times \mathfrak{B}_q) \cdot \bar{\Lambda}, \Delta_{\mathfrak{C}}].$$

Let (c, c') be a generic point of D over k. Put $k_1 = k(c, c')$. As it is well-known, we can extend the specialization ring $[(c, c') \xrightarrow{\mathfrak{o}} (\gamma, \gamma)]$ to a valuation ring $\mathfrak{o}_1$ in k_1.[9''] Since $\dim_k (c, c') = \dim_\kappa (\gamma, \gamma)$, $\mathfrak{o}_1$ is discrete. Put

$$L^{N-n} = (F_i(X - c') = 0 \ (1 \leqq i \leqq n)), \quad \mathfrak{L}^{N-n} = \bar{F}_i(X - \gamma) = 0 \ (1 \leqq i \leqq n)).$$

Then, since $\mathfrak{L}$ is transversal to the tangent linear variety to U at (γ), we have, by Theorem 10, $j(U \cdot \mathfrak{L}, (\gamma)) = 1$. If we denote by $(c^{(1)}), \cdots, (c^{(m)})$ the component of $U \cap L$ such that $\overline{(c^{(i)})} = (\gamma)$ (with respect to a prolongation of $\mathfrak{o}_1$ in an extension of k_1), we have, by Theorem 9

$$1 = \sum_{i=1}^{m} j(U \cdot L, (c^{(i)})) \mu((c^{(i)}), (\gamma));$$

so we have $m = 1$ and $j(U \cdot L, (c^{(1)})) = 1$. As the point (c, c') is in Λ, (c) and (c') are both contained in L; the coordinates of (c) and (c') are all in $\mathfrak{o}_1$; and $\overline{(c)} = \overline{(c')} = (\gamma)$. So we must have $(c) = (c')$; this shows that D is the diagonal Δ_C of a variety C which is the locus of (c) over k. It can be easily verified that C is a proper component of $A \cap B$ on U. For every D_ν, we have a variety C_ν such that $D_\nu = \Delta_{C_\nu}$. Conversely, if we have a component C' of $A \cap B$ such that $\overline{C'} \supset \mathfrak{C}$, then it is easily proved that $\Delta_{C'}$ is

---

[9''] See Theorem 1 and the remark below that theorem.

a component of $\Lambda \cap (A \times B)$; so C' coincides with one of the C_ν. For every ν, the set of linear forms $F_i(X)$ $(1 \leqq i \leqq n)$ is a uniformizing set of linear forms for U along C_ν since $\bar{C}_\nu \supset \mathfrak{C}$ and since the set $\bar{F}_i(X)$ $(1 \leqq i \leqq n)$ is that for U along $\mathfrak{C}$. So we have $i(A \cdot B, C_\nu; U) = j[(A \times B) \cdot \Lambda, \Delta_{C_\nu}]$. By Theorem 8, we have $\mu(A \times B, \mathfrak{A}_p \times \mathfrak{B}_q) = \mu(A, \mathfrak{A}_p)\mu(B, \mathfrak{B}_q)$. Therefore our theorem will be proved if we show that $\mu(\Delta_{C_\nu}, \Delta_{\mathfrak{C}}) = \mu(C_\nu, \mathfrak{C})$. But this is easily proved by means of Proposition 21.[10]

Now, we shall prove the birational invariance of our multiplicity. Let U be a variety in $S^n \times S^m$ defined over k and U' the projection of U on S^n. Let $(x) \times (y)$ be a generic point of U over k and (α) a point in $\overline{U'}$. We say that U is *finite over* (α) if (y) is finite over $(x) \xrightarrow{\mathfrak{o}} (\alpha)$. Also we say that the projection from U to U' is *regular at* (α) if all the y_j are contained in the specialization ring $[(x) \xrightarrow{\mathfrak{o}} (\alpha)]$. Let k' be an extension of k and $\mathfrak{o}'$ a prolongation of $\mathfrak{o}$ in k'. Then by Proposition 6 of Section 1, U is finite over (α) if and only if U is finite over (α) with respect to $\mathfrak{o}'$. Similarly the projection from U to U' is regular at (α) if and only if that projection is regular at (α) with respect to $\mathfrak{o}'$. Let (β) be a specialization of (α) over κ; then the point (β) is a point in $\overline{U'}$. It is easy to see that if U is finite over (β), U is also finite over (α), and that if the projection from U to U' is regular at (β), it is also regular at (α). So we say that U is finite over a variety in $\overline{U'}$ if U is finite over some point in that variety, and also we say that the projection from U to U' is regular along a variety in $\overline{U'}$ if that projection is regular at some point on that variety.

THEOREM 12. *Let U be a variety in $S^n \times S^m$ defined over k, having the same dimension r as its projection U' on S^n. Suppose that there exists a component $\mathfrak{U}'$ of $\overline{U'}$ such that U is finite over $\mathfrak{U}'$. Then there exists a component of $\bar{U}$ whose projection on $\mathfrak{S}^n$ is $\mathfrak{U}'$. If we denote by the $\mathfrak{U}_\nu$ all such components of $\bar{U}$, then we have*

$$[U : U']\mu(U', \mathfrak{U}') = \sum_\nu \mu(U, \mathfrak{U}_\nu)[\mathfrak{U}_\nu : \mathfrak{U}']$$

Proof. Let $L^{n-r} = (\sum t_{ij}X_j - t_i = 0 \ (1 \leqq i \leqq r))$ be a generic linear variety in S^n over k, and $\mathfrak{L}^{n-r} = (\sum \tau_{ij}X_j - \tau_i = 0 \ (1 \leqq i \leqq r))$ a generic linear variety in $\mathfrak{S}^n$ over κ. Then we have a point (x) in $U' \cap L$ and a point (ξ) in $\mathfrak{U}' \cap \mathfrak{L}$. There exists a point (y) in S^m such that (x, y) is a generic point of U over k. By our assumption, (y) is finite over $(x) \xrightarrow{\mathfrak{o}} (\xi)$. Let (η) be a specialization of (y) over $(x) \xrightarrow{\mathfrak{o}} (\xi)$. Then the locus $\mathfrak{U}$ of

---

[10] Let $M = (\sum t_{ij}X_j - t_i = 0 \ (1 \leqq i \leqq r + s - n))$ be a generic linear variety in S^N over k and $\mathfrak{M} = (\sum \tau_{ij}X_j - \tau_i = 0 \ (1 \leqq i \leqq r + s - n))$ a generic linear variety in $\mathfrak{S}^N$ over κ. Apply Proposition 21 to Δ_{C_ν}, $\Delta_{\mathfrak{C}}$, $M \times S^N$ and $\mathfrak{M} \times \mathfrak{S}^N$.

(ξ, η) over $\bar{\kappa}$ is contained in U and has the projection on $\mathfrak{S}^n$. If we denote by $(\eta^{(\nu)})$ all the specializations of (η) over $(x) \xrightarrow{\mathfrak{o}} (\xi)$, then, by Theorem 3 of Section 2, we have

$$\sum \mu_\nu = \mu(U', \mathfrak{U}')[k(t, x, y) : k(t, x)],$$

where μ_ν is the multiplicity of $(\xi, \eta^{(\nu)})$ as a specialization of (x, y) over $(t) \xrightarrow{\mathfrak{o}} (\tau)$. It is obvious that $[U : U'] = [k(t, x, y) : k(t, x)]$. Applying Proposition 21 to U, $\mathfrak{U}_\nu$, $L \times S^m$ and $\mathfrak{L} \times \mathfrak{S}^m$, we have

$$\mu_\nu = \mu(U, \mathfrak{U}_\nu)[\kappa(\tau, \xi, \eta^{(\nu)}) : \kappa(\tau)]_i[\mathfrak{U}_\nu : \kappa]_i^{-1},$$

if we denote by $\mathfrak{U}_\nu$ the locus of $(\xi, \eta^{(\nu)})$ over $\bar{\kappa}$. We have $\mathfrak{U}_\nu = \mathfrak{U}_\lambda$ if and only if $(\eta^{(\nu)})$ and $(\eta^{(\lambda)})$ are conjugates of each other over $\bar{\kappa}(\xi)$. Hence we have

$$\begin{aligned}[U : U']\mu(U', \mathfrak{U}')& \\ &= \sum{}' \mu(U, \mathfrak{U}_\nu)[\bar{\kappa}(\xi, \eta^{(\nu)}) : \bar{\kappa}(\xi)]_s[\kappa(\tau, \xi, \eta^{(\nu)}) : \kappa(\tau)]_i[\mathfrak{U}_\nu : \kappa]_i^{-1}.\end{aligned}$$

where the sum is extended over all the components $\mathfrak{U}_\nu$ of $\bar{U}$ having $\mathfrak{U}'$ as their projections on $\mathfrak{S}^n$. This proves our theorem.

The following theorem is a translation of Theorem 15 of [WF] Chapter IV, Section 7 to our case, and can be proved similarly.

THEOREM 13. *Let U be a variety in $S^n \times S^m$ defined over k with the projection U' on S^n. Let $\mathfrak{Z}'$ be a variety in $\overline{U'}$ such that the projection from U to U' is regular along $\mathfrak{Z}'$. Then there is one and only one variety $\mathfrak{Z}$ in $\bar{U}$ with the projection $\mathfrak{Z}'$ on $\mathfrak{S}^n$; $\mathfrak{Z}$ is simple on U if and only if $\mathfrak{Z}'$ is simple on U; and we have $[\mathfrak{Z} : \mathfrak{Z}'] = 1$.*

Let T be a birational correspondence between two varieties V and W, defined over k. Let $\mathfrak{Z}_V$ be a variety in V such that the projection from T to V is regular along $\mathfrak{Z}_V$. Then, by Theorem 13, there is one and only one variety $\mathfrak{Z}$ in $\bar{T}$ with the projection $\mathfrak{Z}_V$ on $\bar{V}$. If the projection from T to W is also regular along the projection $\mathfrak{Z}_W$ of $\mathfrak{Z}$ on W, then we say that T is *biregular along* $\mathfrak{Z}_V$ and *along* $\mathfrak{Z}_W$, and that $\mathfrak{Z}_V$ and $\mathfrak{Z}_W$ are *regularly corresponding varieties in $\bar{V}$ and in $\bar{W}$ by T.*

THEOREM 14. *Let T be a birational correspondence between two varieties V and V', defined over k. Let A^r be a prime rational cycle over k, which is contained in V, and $\mathfrak{A}$ a prime component of $\rho(A)$ over κ. Suppose that T is biregular along a component of $\mathfrak{A}$ and denote by A' and $\mathfrak{A}'$ the cycles respectively corresponding to A and $\mathfrak{A}$ by T. Then $\mathfrak{A}'$ is a prime component of $\rho(A')$ over κ; and we have $\mu(A, \mathfrak{A}) = \mu(A', \mathfrak{A}')$.*

This follows immediately from Theorem 12, Theorem 13 and from Proposition 20.

Now we shall prove the converse of Theorem 10.

THEOREM 15. *Let* V^r *be a variety in* S^n *defined over* k *and* (ξ) *a point in* $\bar{V}$. *Then* (ξ) *is simple on* V *if and only if* (ξ) *is contained in only one component* $\mathfrak{B}$ *of* $\bar{V}$, $\mu(V,\mathfrak{B})=1$ *and* (ξ) *is simple on* $\mathfrak{B}$. *Moreover, if that is so, and if* (x) *is a generic point of* V *over* k, *then the specialization ring* $[(x)\xrightarrow{\mathfrak{o}}(\xi)]$ *is integrally closed.*

Proof. We have only to prove the direct part. Let the t_{ij} and the s_j $(1\leqq i\leqq r, 1\leqq j\leqq n)$ be $(r+1)n$ independent variables over $k(x)$, let the τ_{ij} and the σ_{ij} be $(r+1)n$ independent variables over $\kappa(\xi)$. Put $y_i=\sum t_{ij}x_j$ and $\eta_i=\sum\tau_{ij}\xi_j$ for $1\leqq i\leqq r$, $z=\sum s_jx_j$ and $\zeta=\sum\sigma_j\xi_j$. Denote by $\mathfrak{L}$ the linear variety $(\sum\tau_{ij}X_j-\eta_i=0\ (1\leqq i\leqq r))$. Since $\mathfrak{L}$ is transversal to the tangent linear variety to $\mathfrak{B}$ at (ξ), we have $j(\mathfrak{B}\cdot\mathfrak{L},(\xi))=1$. By Proposition 9 of Section 1, we can find an extension k_1 of k and a prolongation $\mathfrak{o}_1$ of $\mathfrak{o}$ in k_1 such that the residue field of $\mathfrak{o}_1$ is $\kappa(\xi,\tau)$. We may assume that $k(x,t)$ and k_1 are independent over k. With respect to $\mathfrak{o}_1$, our assumption gives

$$j(V\cdot\mathfrak{L},(\xi))=\mu(V,\mathfrak{B})j(\mathfrak{B}\cdot\mathfrak{L},(\xi))=1,$$

so that (ξ) is a proper specialization of (x) over $(t,y)\xrightarrow{\mathfrak{o}_1}(\tau,\eta)$ of multiplicity 1.[10'] Since $k(x,t)$ and k_1 are linearly disjoint over k, we see that (ξ) is a proper specialization of (x) over $(t,y)\xrightarrow{\mathfrak{o}}(\tau,\eta)$ of multiplicity 1. By Proposition 16 of Section 2, ζ is a proper specialization of z over $(t,s,y)\xrightarrow{\mathfrak{o}}(\tau,\sigma,\eta)$, of multiplicity 1. By the same proposition, we have $k(t,s,y,z)=k(t,s,x)$. Put $k'=k(t,s)$ and $\mathfrak{o}'=[(t,s)\xrightarrow{\mathfrak{o}}(\tau,\sigma)]$. By Proposition 17 of Section 2, the specialization ring $[(y,z)\xrightarrow{\mathfrak{o}'}(\eta,\zeta)]$ is integrally closed, so it contains all the x_j since (x) is finite over this specialization ring by Proposition 16 of Section 2. Let Z be the locus of (x,y,z) over k' and W the locus of (y,z) over k'. Then Z is a birational correspondence between V and W; and the above results shows that (ξ) and (η,ζ) are regularly corresponding points by Z. By Proposition 17 of Section 2, (η,ζ) is simple on W; from this and Theorem 13, follows that (ξ) is simple on V. Moreover we have $[(x)\xrightarrow{\mathfrak{o}'}(\xi)]=[(y,z)\xrightarrow{\mathfrak{o}'}(\eta,\zeta)]$. By this and Proposition 7 of Section 1, $[(x)\xrightarrow{\mathfrak{o}}(\xi)]$ is integrally closed.

COROLLARY 1. *Let* V *be a variety defined over* k *and* $\mathfrak{B}$ *a component of* $\bar{V}$. *Then* $\mathfrak{B}$ *is simple on* V *if and only if* $\mu(V,\mathfrak{B})=1$.

[10'] Note that the linear variety $(\sum t_{ij}X_j-y_i=0\ (1\leqq i\leqq r))$ is generic over k and see the remark below the proof of Theorem 9.

This is an immediate consequence of the above theorem and the corollary of Theorem 10.

COROLLARY 2. *Let U be a variety defined over k and $\mathfrak{U}$ a component of $\bar{U}$; let (x) and (ξ) be respectively a generic point of U over k and a generic point of $\mathfrak{U}$ over $\bar{\kappa}$; and denote by $\mathfrak{R}$ the specialization ring $[(x) \xrightarrow{\mathfrak{o}} (\xi)]$. If $\mathfrak{U}$ is simple on U, then the maximal ideal of $\mathfrak{R}$ is the ideal $\mathfrak{p}\mathfrak{R}$.*

Proof. Let the t_{ij} $(1 \leqq i \leqq r+1, 1 \leqq j \leqq n)$ be $(r+1)n$ independent variables over $k(x)$ and the τ_{ij} $(1 \leqq i \leqq r+1, 1 \leqq j \leqq n)$ be $(r+1)n$ independent variables over $\kappa(\xi)$. By Proposition 9 of Section 1, we can find a finite algebraic extension k_1 of k and a prolongation $\mathfrak{o}_1$ of $\mathfrak{o}$ in k_1 such that $\mathfrak{U}$ is defined over the residue field κ_1 of $\mathfrak{o}_1$. We may assume that the maximal ideal $\mathfrak{p}_1$ of $\mathfrak{o}_1$ is the ideal $\mathfrak{p}\mathfrak{o}_1$.[10''] Put $\mathfrak{o}' = [(t) \xrightarrow{\mathfrak{o}_1} (\tau)]$, $\mathfrak{p}' = \mathfrak{p}_1\mathfrak{o}'$, $\kappa' = \kappa_1(\tau)$, $y_i = \sum t_{ij}x_j$ for $1 \leqq i \leqq r+1$ and $\eta_i = \sum \tau_{ij}\xi_j$ for $1 \leqq i \leqq r+1$. Then as in the proof of Theorem 15, we have $[(x) \xrightarrow{\mathfrak{o}'} (\xi)] = [(y) \xrightarrow{\mathfrak{o}'} (\eta)]$. Moreover, if we denote by V the locus of (y) over $k(t)$, then (η) is a point in $\bar{V}$ and it is simple on V. Let $F(Y) = 0$ be an irreducible equation for (y) over $k(t)$ and $\bar{F}_0(Y) = 0$ an irreducible equation for (η) over κ'. We may assume that $F(Y)$ is a primitive polynomial in $\mathfrak{o}'[Y]$ and that $\bar{F}_0(Y)$ is the class of a ploynomial $F_0(Y)$ in $\mathfrak{o}'[Y]$ modulo $\mathfrak{p}'$. We have $\bar{F}(Y) = \bar{F}_0(Y)\bar{F}_1(Y)$ where $\bar{F}_1(Y)$ is a polynomial in $\kappa'[Y]$. Since (η) is simple on V, we have $\partial\bar{F}/\partial Y_j(\eta) \neq 0$ for some j; so that we have $\bar{F}_1(\eta) \neq 0$. Let z be a quantity in $k(x)$ contained in the maximal ideal of $[(x) \xrightarrow{\mathfrak{o}} (\xi)]$. Then we have an expression $z = P(y)/Q(y)$ where $P(Y)$ and $Q(Y)$ are polynomials in $\mathfrak{o}'[Y]$, $\bar{P}(\eta) = 0$ and $\bar{Q}(\eta) \neq 0$. Since $\bar{F}_0(Y) = 0$ is an irreducible equation for (η) over κ', we have $\bar{P}(Y) = \bar{F}_0(Y)\bar{G}(Y)$ where $\bar{G}(Y)$ is a polynomial in $\kappa'[Y]$. Then we have $\bar{P}(Y)\bar{F}_1(Y) = \bar{F}(Y)\bar{G}(Y)$. Let $F_1(Y)$ and $G(Y)$ be polynomials in $\mathfrak{o}'[Y]$ such that $\bar{F}_1(Y)$ and $\bar{G}(Y)$ are respectively the class of $F_1(Y)$ and the class of $G(Y)$ modulo $\mathfrak{p}'$. Then we have

$$P(Y)F_1(Y) = F(Y)G(Y) + \pi H(Y),$$

where $H(Y)$ is a polynomial in $\mathfrak{o}'[Y]$ and π is a prime element of $\mathfrak{o}$. Substituting (y) for (Y), we have $P(y)F_1(y) = \pi H(y)$, so we have $z/\pi = H(y)/[Q(y)F_1(y)]$. Since $\bar{Q}(\eta)\bar{F}_1(\eta) \neq 0$, z/π is contained in $[(y) \xrightarrow{\mathfrak{o}'} (\eta)]$, so in $[(x) \xrightarrow{\mathfrak{o}'} (\xi)]$. Then by Proposition 7 of Section 1, z/π is contained in $[(x) \xrightarrow{\mathfrak{o}} (\xi)]$. This proves our corollary.

[10''] This is assured since we can take k_1 and $\mathfrak{o}_1$ in such a way that $[k_1 : k] = [\kappa_1 : \kappa]$ as in the proof of Proposition 9.

4. Reduction of abstract varieties. Let $[V_\alpha ; F_\alpha ; T_{\beta\alpha}]$ be an abstract variety in Weil's sense, defined over k; let, for each α, $\mathfrak{F}_\alpha$ be a bunch in $\bar{V}_\alpha$ which is normally algebraic over κ and which contains $\bar{F}_\alpha$. We call the system $[V_\alpha ; F_\alpha ; \mathfrak{F}_\alpha ; T_{\beta\alpha}]$ *a $\mathfrak{p}$-variety defined over k having $[V_\alpha ; F_\alpha ; T_{\beta\alpha}]$ as its underlying abstract variety* if the following condition is satisfied.

Whenever, for any α and β, $\bar{P}_\alpha$ is a point in $\bar{V}_\alpha - \mathfrak{F}_\alpha$ and $\bar{P}_\beta$ a point in $\bar{V}_\beta - \mathfrak{F}_\beta$, such that $(\bar{P}_\alpha, \bar{P}_\beta)$ is in $\bar{T}_{\beta\alpha}$, then $\bar{P}_\alpha$ and $\bar{P}_\beta$ are regularly corresponding points by $T_{\beta\alpha}$.

We can easily prove the following facts.

i) Let $[V_\alpha ; F_\alpha ; \mathfrak{F}_\alpha ; T_{\beta\alpha}]$ be a $\mathfrak{p}$-variety defined over k and

$$[U_{\alpha_\lambda} ; F_{\alpha_\lambda} \cap U_{\alpha_\lambda} ; R_{\mu\lambda}]$$

a subvariety of the abstract variety $[V_\alpha ; F_\alpha ; T_{\beta\alpha}]$. Then the system

$$[U_{\alpha_\lambda} ; F_{\alpha_\lambda} \cap U_{\alpha_\lambda} ; \mathfrak{F}_{\alpha_\lambda} \cap \bar{U}_{\alpha_\lambda} ; R_{\mu\lambda}]$$

defines a $\mathfrak{p}$-variety which will be called a *subvariety* of the $\mathfrak{p}$-variety

$$[V_\alpha ; F_\alpha ; \mathfrak{F}_\alpha ; T_{\beta\alpha}].$$

ii) Let $[V_\alpha ; F_\alpha ; \mathfrak{F}_\alpha ; T_{\beta\alpha}]$ and $[W_\lambda ; G_\lambda ; \mathfrak{G}_\lambda ; S_{\mu\lambda}]$ be two $\mathfrak{p}$-varieties defined over k and $[V_\alpha \times W_\lambda ; H_{\alpha\lambda} ; U_{\beta\mu,\alpha\lambda}]$ the product-variety of the abstract varieties $[V_\alpha ; F_\alpha ; T_{\beta\alpha}]$ and $[W_\lambda ; G_\lambda ; S_{\mu\lambda}]$. Put

$$\mathfrak{H}_{\alpha\lambda} = (\mathfrak{F}_\alpha \times \bar{W}_\lambda) \cup (\bar{V}_\alpha \times \mathfrak{G}_\lambda).$$

Then the system $[V_\alpha \times W_\lambda ; H_{\alpha\lambda} ; \mathfrak{H}_{\alpha\lambda} ; U_{\beta\mu,\alpha\lambda}]$ defines a $\mathfrak{p}$-variety which will be called the *product-variety* of $[V_\alpha ; F_\alpha ; \mathfrak{g}_\alpha ; T_{\beta\alpha}]$ and $[W_\lambda ; G_\lambda ; \mathfrak{G}_\lambda ; S_{\mu\lambda}]$.

iii) The projective space of any dimension defines a $\mathfrak{p}$-variety with the empty $\mathfrak{F}_\alpha$; so that every projective variety defines a uniquely determined $\mathfrak{p}$-variety with the empty $\mathfrak{F}_\alpha$.

Let $[V] = [V_\alpha ; F_\alpha ; \mathfrak{F}_\alpha ; T_{\beta\alpha}]$ be a $\mathfrak{p}$-variety defined over k. Let the M_α be corresponding generic points of the V_α over k by the $T_{\beta\alpha}$ and $(\bar{P}_1, \cdots, \bar{P}_h)$ be any specialization of $(M_1, \cdots, M_h)$ over $\mathfrak{o}$. Similarly as in [WF] p. 168, by a *full set of representatives attached to* $[\bar{V}]$, we understand the set $(\bar{P}_{\alpha_1}, \cdots, \bar{P}_{\alpha_t})$ of all the $\bar{P}_\alpha$ which is finite and not in $\mathfrak{F}_\alpha$. Also we say that $[V]$ is $\mathfrak{p}$-*complete* if no full set of representatives attached to $[\bar{V}]$ is empty. The following facts are easily proved; i) the underlying abstract variety of a $\mathfrak{p}$-complete $\mathfrak{p}$-variety is complete; ii) every subvariety of a $\mathfrak{p}$-complete $\mathfrak{p}$-variety is $\mathfrak{p}$-complete; iii) the product of $\mathfrak{p}$-complete $\mathfrak{p}$-varieties is $\mathfrak{p}$-complete; iv) every projective variety defines a $\mathfrak{p}$-complete $\mathfrak{p}$-variety.

Let $[V]$ be the same as above, and let $[\mathfrak{W}] = [\mathfrak{W}_\lambda; \mathfrak{G}_\lambda; \mathfrak{S}_{\mu\lambda}]$ be an abstract variety defined over an extension field κ' of κ. We say that $[\mathfrak{W}]$ is a *variety in* $[\bar{V}]$ if the following conditions are satisfied: i) there exists a full set of representatives $(\bar{P}_{\alpha_1}, \cdots, P_{\alpha_l})$ attached to $[\bar{V}]$ such that, for every λ, $\mathfrak{W}_\lambda$ is the locus of $\bar{P}_{\alpha_\lambda}$ over κ'; ii) $\mathfrak{G}_\lambda = \mathfrak{W}_\lambda \cap \mathfrak{F}_{\alpha_\lambda}$; iii) $\mathfrak{S}_{\mu\lambda}$ is the variety in $\bar{T}_{\alpha_\mu \alpha_\lambda}$ with the projections $\mathfrak{W}_\lambda$ on $\bar{V}_{\alpha_\lambda}$ and $\mathfrak{W}_\mu$ on $\bar{V}_{\alpha_\mu}$. We call $\mathfrak{W}_\lambda$ the *representative of* $[\mathfrak{W}]$ *in* V_{α_λ}. We can easily prove that for any variety $\mathfrak{W}_\alpha$ in $\bar{V}_\alpha$ which is not contained in $\mathfrak{F}_\alpha$ there is one and only one variety $[\mathfrak{W}]$ in $[\bar{V}]$ such that $\mathfrak{W}_\alpha$ is a representative of $[\mathfrak{W}]$ in V_α. By a *point in* $[\bar{V}]$ we understand a zero-dimensional variety in $[\bar{V}]$. It is obvious that all the representatives of a point in $[\bar{V}]$ forms a full set of representatives attached to $[\bar{V}]$ and conversely. Let $[P]$ be a point in $[V]$ and $[\mathfrak{P}]$ a point in $[\bar{V}]$. We say that $[\mathfrak{P}]$ is a *specialization of* $[P]$ *over* $\mathfrak{o}$ (on $[V]$) if $[P]$ and $[\mathfrak{P}]$ have in some representative V_α of $[V]$, representatives P_α and $\mathfrak{P}_\alpha$ such that $\mathfrak{P}_\alpha$ is a specialization of P_α over $\mathfrak{o}$. We say that a variety $[\mathfrak{W}]$ in $[\bar{V}]$ is *simple on* $[\bar{V}]$ if a representative $\mathfrak{W}_\alpha$ of $[\mathfrak{W}]$ is simple on V_α. We call a finite set of varieties in $[\bar{V}]$ a *bunch in* $[\bar{V}]$; we may identify a bunch with the point-set in $[\bar{V}]$ attached to that bunch. Also, we say that a variety $[\mathfrak{W}]$ is a *component of* a bunch in $[\bar{V}]$ if $[\mathfrak{W}]$ is maximal in that bunch. These definitions are analogies of the definitions in [WF] Chapter VII.

Now our theory of the reduction with respect to $\mathfrak{p}$ can be extended to abstract varieties on which $\mathfrak{p}$-varieties are defined.

Let $[V]$ be a $\mathfrak{p}$-variety defined over k. Let $[A]^r$ be a prime rational $[V]$-cycle over k (which means a prime rational cycle on the underlying abstract variety of $[V]$). We denote by $[\bar{A}]$ the set of all the points in $[\bar{V}]$ which are specializations of some points in $[A]$, over $\mathfrak{o}$. Now we may consider the notation $[\bar{V}]$ in this sense. It is easy to see that the set $[\bar{A}]$ defines a bunch in $[\bar{V}]$. If $[\bar{A}]$ is not empty, then, by Proposition 19 of Section 3, every component of $[\bar{A}]$ is of dimension r. Let $[\mathfrak{A}]_1$ be a component of $[\bar{A}]$. By our definition of $[\bar{A}]$, there exists a representative V_α in which $[A]$ and $[\mathfrak{A}]_1$ have representatives A_α and $\mathfrak{A}_{1\alpha}$.[10] By Theorem 14, of Section 3, the multiplicity $\mu(A_\alpha, \mathfrak{A}_{1\alpha})$ is independent of the choice of representation V_α, A_α, $\mathfrak{A}_{1\alpha}$ for $[V]$, $[A]$ and $[\mathfrak{A}]_1$. We shall denote it by $\mu([A], [\mathfrak{A}]_1)$. Now the results in Section 3 are all translated to the corresponding results on abstract varieties. We shall state them in the following theorems for varieties with a property which we shall call $\mathfrak{p}$-simplicity. We shall say that a $\mathfrak{p}$-variety $[V]$ is *$\mathfrak{p}$-simple* if there exists only one component $[\mathfrak{V}]$ in $[\bar{V}]$ and if $\mu([V], [\mathfrak{V}]) = 1$; and we shall call this $[\mathfrak{V}]$ the *abstract variety obtained from $[V]$ by the reduction with respect to $\mathfrak{p}$.*

---

[10] We shall use the term "representative" also for a prime rational cycle.

THEOREM 16. *Let $[V]$ be a $\mathfrak{p}$-simple $\mathfrak{p}$-variety defined over k, then the abstract variety $[\mathfrak{V}]$ obtained from $[V]$ by the reduction with respect to $\mathfrak{p}$ is defined over κ; and a point on $[\mathfrak{V}]$ is simple on $[\mathfrak{V}]$ if and only if it is simple on $[V]$.*

Proof. As for the first assertion, it is sufficient to prove that every representative of $[\mathfrak{V}]$ is defined over κ. Let V_α be a representative of $[V]$ and $\mathfrak{V}_\alpha$ a representative of $[\mathfrak{V}]$ in V_α. By Corollary 1 of Theorem 15 of Section 3 and by corollary of Theorem 10 of Section 3, we have $[\mathfrak{V}_\alpha : \kappa]_i = 1$. This shows $[\kappa' : \kappa]_i = 1$ if we denote by κ' the smallest field of definition for $\mathfrak{V}_\alpha$ which contains κ. Since the bunch $\mathfrak{F}_\alpha$ in $\bar{V}_\alpha$ is normally algebraic over κ, every conjugate of $\mathfrak{V}_\alpha$ over κ is not contained in $\mathfrak{F}_\alpha$; so that, by our assumption, there is no conjugate of $\mathfrak{V}_\alpha$ over κ other than $\mathfrak{V}_\alpha$ itself; so we have $[\kappa' : \kappa]_s = 1$. Hence we have $\kappa' = \kappa$. The remaining part of the theorem follows immediately from Theorem 15 of Section 3.

PROPOSITION 23. *If $[U]$ and $[V]$ are $\mathfrak{p}$-simple $\mathfrak{p}$-varieties defined over k, then the product-variety $[U] \times [V]$ is also $\mathfrak{p}$-simple.*

This is easily proved since the result corresponding to Proposition 5 of [WF] Chapter VII, Section 3 also holds in our case.

Let $[V]$ be a $\mathfrak{p}$-simple variety defined over k and $[\mathfrak{V}]$ the abstract variety obtained from $[V]$ by the reduction with respect to $\mathfrak{p}$. Then every "variety in $[\bar{V}]$" is a subvariety of $[\mathfrak{V}]$. Thus we are able to define the reduction of $[V]$-cycles. Let $[A]^r$ be a prime rational $[V]$-cycle over k and $[\mathfrak{A}]_1{}^r$ a prime rational $[\mathfrak{V}]$-cycle over κ with a component $[\mathfrak{A}]_{11}$ which is a component of $[\bar{A}]$. We denote by $\mu([A], [\mathfrak{A}]_1)$ the number

$$\mu([A], [\mathfrak{A}]_{11})/[[\mathfrak{A}]_{11} : \kappa]_i,$$

and define the $[\mathfrak{V}]$-cycle $\rho([A])$ as follows:

$$\rho([A]) = \sum_\nu \mu([A], [\mathfrak{A}]_\nu)[\mathfrak{A}]_\nu,$$

where the sum is taken over all the prime rational $[\mathfrak{V}]$-cycles $[\mathfrak{A}]_\nu$ over κ with the components which are components of $[\bar{A}]$. Even if $[\bar{A}]$ is not empty the cycle $\rho([A])$ vanishes when every component of $[\bar{A}]$ is not simple on $[\mathfrak{V}]$. Let $[X]$ be a rational $[V]$-cycle over k; we have an expression $[X] = \sum_\lambda a_\lambda [A]_\lambda$ where, for every λ, $[A]_\lambda$ is a prime rational $[V]$-cycle over k. We put

$\rho([X]) = \sum_{\lambda} a_{\lambda}\rho([A]_{\lambda})$ and call it the *cycle obtained from* $[X]$ *by the reduction with respect to* $\mathfrak{p}$. It is obvious that ρ is a linear mapping of the group of rational $[V]$-cycles over k into the group of rational $[\mathfrak{V}]$-cycles over κ. Similarly as in the affine space, $\rho([X])$ is invariant under the extension of the basic field k and the prolongation of $\mathfrak{o}$. It should be mentioned that the mapping ρ is defined with reference to the ambient variety $[V]$. But we shall use the same notation ρ for mappings of cycles on different ambient varieties, since there will be no confusion.

THEOREM 17. *Let* $[V]$ *be a* $\mathfrak{p}$*-simple* $\mathfrak{p}$*-variety defined over* k. *Let* $[X]^r$ *and* $[Y]^s$ *be positive rational* $[V]$*-cycles over* k *such that the intersection-products* $[X]\cdot[Y]$ *and* $\rho([X])\cdot\rho([Y])$ *are both defined. Then we have* $\rho([X]\cdot[Y]) = \rho([X])\cdot\rho([Y])$.

Proof. By linearity it is sufficient to prove our theorem in case where $[X]$ and $[Y]$ are prime rational $[V]$-cycles $[A]^r$ and $[B]^s$ over k, respectively. Moreover we may assume that $[A]$ and $[B]$ are both varieties defined over k. For, we may take a suitable extension of k and a prolongation of $\mathfrak{o}$ instead of k and $\mathfrak{o}$, since ρ is invariant under such a change of the basic field and ring. We may also assume that every component of $[A] \cap [B]$ is defined over k. Now denote by $[\mathfrak{V}]$ the abstract variety obtained from $[V]$ by reduction. Let $[C]$ be a proper component of $[A] \cap [B]$ on $[V]$ and $[\mathfrak{C}]$ a component of $[\bar{C}]$. If $[\mathfrak{C}]$ is simple on $[\mathfrak{V}]$, it is a component of $\rho([A])\cdot\rho([B])$. To see this let $[\mathfrak{C}]'$ be a component of $[\bar{A}] \cap [\bar{B}]$ containing $[\mathfrak{C}]$; then $[\mathfrak{C}]'$ is simple on $[\mathfrak{V}]$. By our assumption that $\rho([A])\cdot\rho([B])$ is defined on $[\mathfrak{V}]$, we have $\dim [\mathfrak{C}]' = r + s - n$ where n is the dimension of $[V]$ and of $[\mathfrak{V}]$; so that we have $\dim [\mathfrak{C}]' = \dim [C] = \dim [\mathfrak{C}]$. Hence $[\mathfrak{C}]$ is a component of $\rho([A])\cdot\rho([B])$. Conversely, let $[\mathfrak{C}]$ be a component of $\rho([A])\cdot\rho([B])$. Then it is a component of $[\bar{A}] \cap [\bar{B}]$ of dimension $r + s - n$ and is simple on $[\mathfrak{V}]$. By Theorem 16, $[\mathfrak{C}]$ is also simple on $[V]$; then, by Theorem 11 of Section 3, there exists a component $[C]$ of $[A] \cap [B]$ such that $[\mathfrak{C}]$ is a component of $[\bar{C}]$. Hence we see that $[\mathfrak{C}]$ is a component of $\rho([A]\cdot[B])$. Thus we have proved that a variety on $[\mathfrak{V}]$ is a component of $\rho([A]\cdot[B])$ if and only if it is a component of $\rho([A])\cdot\rho([B])$. Now our theorem is an immediate consequence of the formula in Theorem 11 of Section 3.

THEOREM 18. *Let* $[U]$ *and* $[V]$ *be* $\mathfrak{p}$*-simple* $\mathfrak{p}$*-varieties defined over* k. *If* $[X]$ *is a rational* $[U]$*-cycle over* k *and* $[Y]$ *is a rational* $[V]$*-cycle over* k, *then we have* $\rho([X] \times [Y]) = \rho([X]) \times \rho([Y])$.

This follows immediately from Theorem 8 of Section 3 and from our definition of ρ.

THEOREM 19. *Let* $[U]$ *and* $[V]$ *be* $\mathfrak{p}$*-simple* $\mathfrak{p}$*-varieties defined over* k*; and let* $[\mathfrak{U}]$ *and* $[\mathfrak{V}]$ *be respectively the abstract varieties obtained from* $[U]$ *and* $[V]$ *by the reduction with respect to* $\mathfrak{p}$*. Suppose that* $[V]$ *is* $\mathfrak{p}$*-complete and that* $[\mathfrak{V}]$ *has no multiple point. If* $[X]$ *is a rational* $[U]\times[V]$*-cycle over* k*, then we have* $\rho\{\mathrm{pr}_{[U]}([X])\}=\mathrm{pr}_{[\mathfrak{U}]}\{\rho([X])\}$.

Proof. By linearity, it is enough to prove our theorem in case where $[X]$ is a prime rational cycle $[A]$. Moreover, as in the proof of Theorem 17, we may assume that $[A]$ is a variety defined over k. Let $[A]'$ be the projection of $[A]$ on $[U]$, and $[P]\times[Q]$ a generic point of $[A]$ over k. Our theorem is obvious when $\dim[A]'<\dim[A]$; so we assume that $[A]'$ has the same dimension r as $[A]$ and put $[[A]:[A]']=d$. Let $[\mathfrak{A}]'$ be a component of $\rho([A]')$ and $[\bar{P}]$ a generic point of $[\mathfrak{A}]'$ over $\bar{\kappa}$. Since $[V]$ is $\mathfrak{p}$-complete, there exists a point $[\bar{Q}]$ on $[\mathfrak{V}]$ which is a specialization of $[Q]$ over $[P]\xrightarrow{\mathfrak{o}}[\bar{P}]$. If we denote by $[\mathfrak{A}]$ the locus of $[\bar{P}]\times[\bar{Q}]$ over $\bar{\kappa}$, then $[\mathfrak{A}]$ is a component of $\rho([A])$ with the projection $[\mathfrak{A}]'$ on $[\mathfrak{U}]$. Thus every component of $\rho\{\mathrm{pr}_{[U]}([A])\}$ is a component of $\mathrm{pr}_{[\mathfrak{U}]}\{\rho([A])\}$. It is easy to see that the converse is also true. Now we may assume that the components of $\rho([A])$ and $\rho([A]')$ are all defined over κ. Our theorem is proved when we show that, for every component $[\mathfrak{A}]'$ of $\rho([A]')$,

$$d\cdot\mu([A]',[\mathfrak{A}]')=\sum_{\sigma}\mu([A],[\mathfrak{A}]_\sigma)\,[[\mathfrak{A}]_\sigma:[\mathfrak{A}]'],$$

where the sum is taken over every component $[\mathfrak{A}]_\sigma$ of $\rho([A])$ with the projection $[\mathfrak{A}]'$ on $[\mathfrak{U}]$. Let U, $\mathfrak{U}$, A', $\mathfrak{A}'$, P, $\bar{P}$ be representations of $[U]$, $[\mathfrak{U}]$, $[A]'$ $[\mathfrak{A}]'$, $[P]$ and $[\bar{P}]$, respectively; and let S^N and $\mathfrak{S}^N$ be respectively the ambient spaces for U and for $\mathfrak{U}$. We may assume that $[U]=U$ and $[\mathfrak{U}]=\mathfrak{U}$. Let $L^{N-r}=(\sum t_{ij}X_j-t_i=0\ \ (1\leqq i\leqq r))$ be a generic linear variety in S^N over k and $\mathfrak{L}^{N-r}=(\sum\tau_{ij}X_j-\tau_i=0\ \ (1\leqq i\leqq r))$ a generic linear variety in $\mathfrak{S}^N$ over κ. We may assume that the point P is in L and that $\bar{P}$ is in $\mathfrak{L}$. Denote by $(P_1,\cdots,P_\delta)$ the complete set of conjugates of P over $k(t)$ and denote by $([Q]_1,\cdots,[Q]_d)$ the complete set of conjugates of $[Q]$ over $k(P)$; then we have the set of δd points $P_\lambda\times[Q]_{\lambda\nu}$ $(1\leqq\lambda\leqq\delta,1\leqq\nu\leqq d)$ as the complete set of conjugates of $P\times[Q]$ over $k(t)$, where $([Q]_{\lambda1},\cdots,[Q]_{\lambda d})$ is the transform of $([Q]_1,\cdots,[Q]_d)$ by some automorphism over $k(t)$ of $\overline{k(t)}$. Let $(\bar{P}_\lambda\times[\bar{Q}]_{\lambda\nu})$ be a specialization of $(P_\lambda\times[Q]_{\lambda\nu})$ over $(t)\xrightarrow{\mathfrak{o}}(\tau)$. By our assumption that $[V]$ is $\mathfrak{p}$-complete, $[\bar{Q}]_{\lambda\nu}$ is a point of $[\mathfrak{V}]$ for every (λ,ν). If $\bar{P}_\lambda$ is a conjugate of $\bar{P}$

over $\kappa(\tau)$, then $\bar{P}_\lambda \times [\bar{Q}]_{\lambda\nu}$ is a generic point of one of the $[\mathfrak{A}]_\sigma$, say $[\mathfrak{A}]_1$, over κ. If $[\mathfrak{A}]_1$ has a representative in $U \times V_\alpha$ then every generic point of $[\mathfrak{A}]_1$ over κ has a representative in $U \times V_\alpha$. Hence we may consider that we are in the affine space; so that by Proposition 21 of Section 3,[11] the number of (λ, ν) such that $\bar{P}_\lambda \times [\bar{Q}]_{\lambda\nu}$ is a generic point of $[\mathfrak{A}]_1$ over k is equal to $\mu([A], [\mathfrak{A}]_1)[\kappa(\bar{P}_\lambda \times [\bar{Q}]_{\lambda\nu}, \tau) : \kappa(\tau)]$. On the other hand, the number of $\bar{P}_\lambda \times [\bar{Q}]_{\lambda\nu}$ such that $\bar{P}_\lambda$ is a conjugate of $\bar{P}$ over $\kappa(\tau)$ is equal to

$$d \cdot \mu([A]', [\mathfrak{A}]')[\kappa(\bar{P}, \tau) : \kappa(\tau)].$$

Hence we have

$$d \cdot \mu([A]', [\mathfrak{A}]')[\kappa(\bar{P}, \tau) : \kappa(\tau)] = \sum_\sigma \mu([A], [\mathfrak{A}]_\sigma)[[\mathfrak{A}]_\sigma : [\mathfrak{A}]'][\kappa(\bar{P}, \tau) : \kappa(\tau)].$$

This completes our proof.

PROPOSITION 24. *Let $[V]$ be a $\mathfrak{p}$-simple $\mathfrak{p}$-variety defined over k and $[\mathfrak{V}]$ the abstract variety obtained from $[V]$ by the reduction with respect to $\mathfrak{p}$. Suppose that $[V]$ is $\mathfrak{p}$-complete and $[\mathfrak{V}]$ has no multiple point. Then for every rational $[V]$-cycle $[X]$ over k, of dimension* 0, *we have*

$$\deg([X]) = \deg\{\rho([X])\}.$$

Proof. As in the proof of Theorem 17, we may assume that every component of $[X]$ is rational over k. Since every rational point over k has a uniquely determined specialization over $\mathfrak{o}$, our proposition is an immediate consequence of our definition of ρ.

PROPOSITION 25. *Let $[L]^n$ be the projective space defined over k and $[X]$ a rational $[L]$-cycle over k. Consider the abstract variety obtained from $[L]$ by the reduction as the projective space $[\mathfrak{L}]^n$ defined over κ. Then we have* $\deg([X]) = \deg\{\rho([X])\}$.

This follows immediately from Theorem 17 and Proposition 24.

Let $[V]$ be a $\mathfrak{p}$-simple $\mathfrak{p}$-variety defined over k and $[\mathfrak{V}]$ the abstract variety obtained from $[V]$ by the reduction with respect to $\mathfrak{p}$. We denote by $\boldsymbol{F}$ the field of algebraic functions on $[V]$ defined over k and by Φ the field of algebraic functions on $[\mathfrak{V}]$ defined over κ. Let $[P]$ be a generic point of $[V]$ over k and $[\bar{P}]$ a generic point of $[\mathfrak{V}]$ over κ. Let ϕ be an element of $\boldsymbol{F}$ which is defined by $z = \phi([P])$. By our assumption and by Theorem 15 of Section 3, the specialization ring $[[P] \xrightarrow{\mathfrak{o}} [\bar{P}]]$ is integrally closed, so that by Proposition 5 of Section 1, it is a discrete valuation ring. Hence z has a uniquely determined specialization ζ over $[P] \xrightarrow{o} [\bar{P}]$; and ζ is a generalized quantity in $\kappa([\bar{P}])$. Denote by $\bar{\phi}$ the generalized function on $[\mathfrak{V}]$ defined by $\zeta = \bar{\phi}([\bar{P}])$ over κ. Thus we obtain a mapping $\phi \to \bar{\phi}$ from $\boldsymbol{F}$ to Φ; it is easy to see that this mapping does not depend on the choice of $[P]$ and $[\bar{P}]$. Therefore, we may call the field Φ the *field obtained from* $\boldsymbol{F}$ *by the reduction with respect to* $\mathfrak{p}$; of course, the field Φ depends upon the variety $[V]$ which is a "model" of $\boldsymbol{F}$.

---

[11] Apply that proposition to $[A]$, $[\mathfrak{A}]_1$, $L \times S^m$ and $\mathfrak{L} \times \mathfrak{S}^m$, where S^m and $\mathfrak{S}^m$ denote the ambient spaces for V_α and $\bar{V}_\alpha$, respectively.

THEOREM 20. *Let* $[V]$ *be a* $\mathfrak{p}$*-simple* $\mathfrak{p}$*-variety defined over* k *and* $[\mathfrak{V}]$ *the abstract variety obtained from* $[V]$ *by the reduction with respect to* $\mathfrak{p}$. *Let* ϕ *be a function on* $[V]$ *defined over* k *such that* $\bar{\phi}$ *is a function on* $[\mathfrak{V}]$ *other than the constant* 0. *Then* $\rho((\phi)) = (\bar{\phi})$.

Proof. Let $[D]$ be the projective straight line defined over k and $[\mathfrak{D}]$ be the projective straight line defined over κ. We may consider $[\mathfrak{D}]$ as the abstract variety obtained from $[D]$ by reduction. Let Γ_ϕ be the graph of ϕ and $\Gamma_{\bar{\phi}}$ the graph of $\bar{\phi}$; then $\Gamma_{\bar{\phi}}$ is a component of $\bar{\Gamma}_\phi$. By Theorem 19, we have

$$\mathrm{pr}_{[\mathfrak{V}]}\{\rho(\Gamma_\phi)\} = \rho\{\mathrm{pr}_{[V]}(\Gamma_\phi)\} = \rho([V]) = [\mathfrak{V}].$$

By this equation we have $\rho(\Gamma_\phi) = \Gamma_{\bar{\phi}} + [\mathfrak{X}] \times [\mathfrak{D}]$, where $[\mathfrak{X}]$ is a $[\mathfrak{V}]$-divisor. Now by the definition of $(\phi)_0$, Theorem 17, Theorem 18, and Theorem 19, we have

$$\begin{aligned}\rho((\phi)_0) &= \rho\{\mathrm{pr}_{[V]}(\Gamma_\phi \cdot [V]_0)\} = \mathrm{pr}_{[\mathfrak{V}]}\{(\Gamma_{\bar{\phi}} + [\mathfrak{X}] \times [\mathfrak{D}]) \cdot ([\mathfrak{V}] \times (0))\} \\ &= \mathrm{pr}_{[\mathfrak{V}]}\{\Gamma_{\bar{\phi}} \cdot [\mathfrak{V}]_0 + [\mathfrak{X}] \times (0)\} = (\bar{\phi})_0 + [\mathfrak{X}].\end{aligned}$$

Similarly, we have $\rho((\phi)_\infty) = (\bar{\phi})_\infty + [\mathfrak{X}]$; so we have $\rho((\phi)) = (\bar{\phi})$.

THEOREM 21. *Let* A^{n-1} *be a prime rational* S^n*-divisor over* k *and* (x) *a generic point of* A *over* k. *Let* $F(X) = 0$ *be an irreducible equation for* (x) *over* k *and suppose that* $F(X)$ *is a primitive polynomial in* $\mathfrak{o}[X]$. *Let* $\bar{F}(X) = \prod_\alpha \bar{G}_\alpha(X)^{\mu_\alpha}$ *be an expression of* $\bar{F}(X)$ *as a product of irreducible polynomials in* $\kappa[X]$. *Then, we have* $\rho(A) = \sum_\alpha \mu_\alpha \mathfrak{A}_\alpha$ *where, for each* α, $\mathfrak{A}_\alpha$ *is the prime rational* $\mathfrak{S}^n$*-divisor over* κ *which has a generic point* $(\xi^{(\alpha)})$ *over* κ *such that* $\bar{G}_\alpha(\xi^{(\alpha)}) = 0$.

Proof. Let $P = (t)$ and $\bar{P} = (\tau)$ be respectively a generic point of S^n over k and a generic point of $\mathfrak{S}^n$ over κ. Let ϕ be the function on S^n, defined over k by $F(t) = \phi(P)$. Then the function $\bar{\phi}$ on $\mathfrak{S}^n$ is defined over κ by $\bar{F}(\tau) = \bar{\phi}(\bar{P})$. By our assumption on $F(X)$, the function $\bar{\phi}$ is not the constant 0; so that by Theorem 8 of [WF] Chapter VIII, Section 3 and by Theorem 20, we have $\rho(A) = \rho((\phi)) = (\bar{\phi})$. Let $\bar{\psi}_\alpha$ be the function on $\mathfrak{S}^n$ defined over κ by $\bar{G}_\alpha(\tau) = \bar{\psi}_\alpha(\bar{P})$, for each α. Then we have $\bar{\phi} = \prod_\alpha \bar{\psi}_\alpha^{\mu_\alpha}$; so we have by Theorem 6 and Theorem 8 of [WF] Chapter VIII, $(\bar{\phi}) = \sum_\alpha \mu_\alpha(\bar{\psi}_\alpha) = \sum_\alpha \mu_\alpha \mathfrak{A}_\alpha$. This proves our theorem.

5. Specialization-theory of cycles. When we consider the reduction of cycles with respect to $\mathfrak{p}$ as the specialization of cycles, our theory in Section 4 is concerned only with the specialization of the cycles which are rational over the basic field. In this section, we shall define a more general concept of the specialization of cycles on abstract varieties.

Let $[V]_1, \cdots, [V]_n$ be $\mathfrak{p}$-simple $\mathfrak{p}$-varieties defined over k and $[\mathfrak{V}]_1, \cdots, [\mathfrak{V}]_n$ be the abstract varieties obtained from $[V]_1, \cdots, [V]_n$ by the reduction with respect to $\mathfrak{p}$, respectively. Let $[X]_1, \cdots, [X]_n$ be respectively cycles on $[V]_1, \cdots, [V]_n$; and let $[\mathfrak{X}]_1, \cdots, [\mathfrak{X}]_n$ be respectively cycles on $[\mathfrak{V}]_1, \cdots, [\mathfrak{V}]_n$. We say that $([\mathfrak{X}]_1, \cdots, [\mathfrak{X}]_n)$ is a *specialization of* $([X]_1, \cdots, [X]_n)$ *over* $\mathfrak{o}$ if there exist an extension k' of k and a prolongation $\mathfrak{o}'$ of $\mathfrak{o}$ in k' such that $[X]_1, \cdots, [X]_n$ are all rational over k' and $[\mathfrak{X}]_1, \cdots, [\mathfrak{X}]_n$ are respectively the cycles obtained from $[X]_1, \cdots, [X]_n$ by the reduction with respect to $\mathfrak{o}'$. We shall denote by $([X]_1, \cdots, [X]_n) \xrightarrow{\mathfrak{o}} ([\mathfrak{X}]_1, \cdots, [\mathfrak{X}]_n)$.

PROPOSITION 26. *Let* (x) *be a set of* n *quantities in* $\boldsymbol{K}$ *and* (ξ) *a specialization of* (x) *over* $\mathfrak{o}$. *Then there exists a prolongation* $\mathfrak{o}'$ *of* $\mathfrak{o}$ *in* $k(x)$ *such that* (ξ) *is a specialization of* (x) *over* $\mathfrak{o}'$.

Proof. By Proposition 9 of Section 1, there exist an extension k_1 of k and a prolongation $\mathfrak{o}_1$ of $\mathfrak{o}$ in k_1 such that the residue field κ_1 of $\mathfrak{o}_1$ is $\kappa(\xi)$. Let (b) be a set of quantities in $\mathfrak{o}_1$ such that $(b) \xrightarrow{\mathfrak{o}_1} (\xi)$. We may assume that $k_1 = k(b)$. Let r be the dimension of (x) over k. Let the t_{ij} and the t_i for $1 \leqq i \leqq r$, $1 \leqq j \leqq n$ be $r(n+1)$ independent variables over k_1; and let the τ_{ij} and the τ_i for $1 \leqq i \leqq r$, $1 \leqq j \leqq n$ be $r(n+1)$ independent variables over κ_1. Put $k_2 = k_1(t_i, t_{ij})$ and $\mathfrak{o}_2 = [(t_i, t_{ij}) \xrightarrow{\mathfrak{o}_1} (\tau_i, \tau_{ij})]$; then $\mathfrak{o}_2$ is a prolongation of $\mathfrak{o}_1$ in k_2. Denote by L^{n-r} the linear variety

$$(\textstyle\sum t_{ij}(X_j - b_j) - \pi t_i = 0 \quad (1 \leqq i \leqq r))$$

in S^n where π is a prime element in $\mathfrak{o}$ and by $\mathfrak{L}^{n-r}$ the linear variety

$$(\textstyle\sum \tau_{ij}(X_j - \xi_j) = 0 \quad (1 \leqq i \leqq r))$$

in $\mathfrak{S}^n$; denote by A^r the prime rational cycle in S^n with the generic point (x) over k. Now we consider the reduction with respect to $\mathfrak{o}_2$. Then we have $\bar{L} = \mathfrak{L}$. The point (ξ) is contained in $\bar{A} \cap \mathfrak{L}$; moreover, since (τ) is a set of independent variables over κ_1, it can be easily proved that (ξ) is a component of $\bar{A} \cap \mathfrak{L}$. By Theorem 9 of Section 2, there exists a point (y) in $A \cap L$ such that (ξ) is a specialization of (y) over $\mathfrak{o}_2$. We extend the specialization ring $[(y) \xrightarrow{\mathfrak{o}_2} (\xi)]$ to a valuation ring $\mathfrak{o}_3$ in $k_3 = k_2(y)$; this valuation ring $\mathfrak{o}_3$ is also discrete as (y) is algebraic over k. Since L is a generic linear variety over k, we have $(x) \overset{k}{\longleftrightarrow} (y)$. We extend this generic specialization to a generic specialization $(x, a, s) \overset{k}{\longleftrightarrow} (y, b, t)$. Put $k_4 = k(x, a, s)$ and denote by $\mathfrak{o}_4$ the prolongation of $\mathfrak{o}$ in k_4 which corresponds to the prolongation $\mathfrak{o}_3$ in k_3 by the specialization $(x, a, s) \overset{k}{\longleftrightarrow} (y, b, t)$. Put $\mathfrak{o}' = k(x) \cap \mathfrak{o}_4$; then $\mathfrak{o}'$ is a discrete valuation ring in $k(x)$ and it defines a prolongation of $\mathfrak{o}$ in $k(x)$ such that $(x) \xrightarrow{\mathfrak{o}} (\xi)$; so our proposition is proved.

By this proposition, our definition of the specialization of cycles does not conflict with the definition of the specialization of points. But it should be remarked that singular points or a pseudo-point are considered as the cycle 0.

PROPOSITION 27. *Let* $[U]_1, \cdots, [U]_n$, $[V]_1, \cdots, [V]_m$ *be* $\mathfrak{p}$*-simple* $\mathfrak{p}$*-varieties defined over* k; *and let* $[X]_1, \cdots, [X]_n$, $[Y]_1, \cdots, [Y]_m$ *be respectively cycles on* $[U]_1, \cdots, [U]_n$, $[V]_1, \cdots, [V]_m$. *Every specialization of* $([X]_1, \cdots, [X]_n)$ *over* $\mathfrak{o}$ *can be extended to a specialization of* $([X]_1, \cdots, [X]_n, [Y]_1, \cdots, [Y]_m)$ *over* $\mathfrak{o}$.

Proof. Let $([\mathfrak{X}]_1, \cdots, [\mathfrak{X}]_n)$ be a specialization of $([X]_1, \cdots, [X]_n)$ over $\mathfrak{o}$. Then, by our definition, there exist an extension k' of k and a prolongation $\mathfrak{o}'$ of $\mathfrak{o}$ in k' such that all the $[X]_i$ are rational over k' and that $\rho'([X]_i) = [\mathfrak{X}]_i$ for $1 \leqq i \leqq n$ where ρ' denotes the reduction-mapping with respect to $\mathfrak{o}'$. Let k'' be an extension of k' such that $[Y]_1, \cdots, [Y]_m$ are all rational over k'' and $\mathfrak{o}''$ a prolongation of $\mathfrak{o}'$ in k''. Put $[\mathfrak{Y}]_i = \rho''([Y]_i)$ for $1 \leqq i \leqq m$ where ρ'' denotes the reduction-mapping with respect to $\mathfrak{o}''$. Since $\rho''([X]_i) = \rho'([X]_i)$, we have

$$([X]_1, \cdots, [X]_n, [Y]_1, \cdots, [Y]_m) \xrightarrow{\mathfrak{o}} ([\mathfrak{X}]_1, \cdots, [\mathfrak{X}]_n, [\mathfrak{Y}]_1, \cdots, [\mathfrak{Y}]_m).$$

This proves our proposition.

Now we can easily prove that the specialization of cycles defined above preserves the operations on cycles.

THEOREM 22. *Let* $[V]$ *be a* $\mathfrak{p}$*-simple* $\mathfrak{p}$*-variety defined over* k, $[X]$ *and* $[Y]$ *two* $[V]$*-cycles of the same dimension and* $([\mathfrak{X}], [\mathfrak{Y}])$ *a specialization* *of* $([X],[Y])$ *over* $\mathfrak{o}$. *Then* $[\mathfrak{X}]+[\mathfrak{Y}]$ *is a uniquely determined specialization of* $[X]+[Y]$ *over* $([X],[Y])\xrightarrow{\mathfrak{o}}([\mathfrak{X}],[\mathfrak{Y}])$.

THEOREM 23. *Let* $[V]$ *be a* $\mathfrak{p}$*-simple* $\mathfrak{p}$*-variety defined over* k, $[X]^r$ *and* $[Y]^s$ *two positive* $[V]$*-cycles and* $([\mathfrak{X}],[\mathfrak{Y}])$ *a specialization of* $([X],[Y])$ *over* $\mathfrak{o}$. *If* $[X]\cdot[Y]$ *and* $[\mathfrak{X}]\cdot[\mathfrak{Y}]$ *are both defined, then* $[\mathfrak{X}]\cdot[\mathfrak{Y}]$ *is a uniquely determined specialization of* $[X]\cdot[Y]$ *over*

$$([X],[Y])\xrightarrow{\mathfrak{o}}([\mathfrak{X}],[\mathfrak{Y}]).$$

THEOREM 24. *Let* $[U]$ *and* $[V]$ *be two* $\mathfrak{p}$*-simple* $\mathfrak{p}$*-varieties defined over* k, $[X]\times[Y]$ *a* $[U]\times[V]$*-cycle and* $([\mathfrak{X}],[\mathfrak{Y}])$ *a specialization of* $([X],[Y])$ *over* $\mathfrak{o}$. *Then,* $[\mathfrak{X}]\times[\mathfrak{Y}]$ *is a uniquely determined specialization of* $[X]\times[Y]$ *over* $([X],[Y])\xrightarrow{\mathfrak{o}}([\mathfrak{X}],[\mathfrak{Y}])$.

THEOREM 25. *Let* $[U]$ *and* $[V]$ *be two* $\mathfrak{p}$*-simple* $\mathfrak{p}$*-varieties defined over* k *and denote by* $[\mathfrak{U}]$ *and* $[\mathfrak{V}]$ *the varieties obtained from* $[U]$ *and* $[V]$ *by the reduction with respect to* $\mathfrak{p}$, *respectively. Suppose that* $[V]$ *is* $\mathfrak{p}$*-complete and that* $[\mathfrak{V}]$ *has no multiple point. Let* $[X]$ *be a* $[U]\times[V]$*-cycle and* $[\mathfrak{X}]$ *a specialization of* $[X]$ *over* $\mathfrak{o}$. *Then* $\mathrm{pr}_{[\mathfrak{U}]}([\mathfrak{X}])$ *is a uniquely determined specialization of* $\mathrm{pr}_{[U]}([X])$ *over* $[X]\xrightarrow{\mathfrak{o}}[\mathfrak{X}]$.

We shall only prove Theorem 23 since the others are proved in the same way.

Proof of Theorem 23. Let $([\mathfrak{X}],[\mathfrak{Y}],[\mathfrak{Z}])$ be a specialization of $([X],[Y],[X]\cdot[Y])$ over $\mathfrak{o}$. By our definition, there exist an extension k' of k and a prolongation $\mathfrak{o}'$ of $\mathfrak{o}$ in k' such that $[X]$, $[Y]$ are rational over k' and $\rho'([X])=[\mathfrak{X}]$, $\rho'(Y)=[\mathfrak{Y}]$, $\rho'([X]\cdot[Y])=[\mathfrak{Z}]$, where ρ' denotes the reduction-mapping defined with respect to $\mathfrak{o}'$. Since $[\mathfrak{X}]\cdot[\mathfrak{Y}]$ is defined by our assumption, we have by Theorem 18 of Section 4, $\rho'([X]\cdot[Y])=[\mathfrak{X}]\cdot[\mathfrak{Y}]$. This proves our theorem.

In case where $k=\mathfrak{o}$, we have a specialization-theory of cycles over a field. In this case, we have $[\mathfrak{U}]=[U]$ (or $[\mathfrak{U}]=[U]^\sigma$ where σ is an isomorphism of k) for an ambient variety $[U]$.

We have to prove that the specialization of cycles thus defined is transitive since this is not so obvious.

PROPOSITION 28. *Let* $[V]$ *be a* $\mathfrak{p}$*-simple* $\mathfrak{p}$*-variety defined over* k, $[X]'$ *a* V*-cycle and* $[\mathfrak{X}]$ *a specialization of* $[X]'$ *over* $\mathfrak{o}$.

i) *If* $[X]'$ *is a specialization of a* $[V]$*-cycle* $[X]$ *over* k, *then* $[\mathfrak{X}]$ *is a specialization of* $[X]$ *over* $\mathfrak{o}$.

ii) *Every specialization of* $[\mathfrak{X}]$ *over* κ *is also a specialization of* $[X]'$ *over* $\mathfrak{o}$.

Proof. Suppose that $[X] \xrightarrow{k} [X]' \xrightarrow{\mathfrak{o}} [\mathfrak{X}]$. Then, by our definition, there exist an extension k_1 of k and a discrete valuation ring $\mathfrak{o}_1$ in k_1 which contains k such that $[X]$ is rational over k_1 and $\rho_1([X]) = [X]'$ where ρ_1 is the reduction-mapping with respect to $\mathfrak{o}_1$; and there exist an extension k_2 of k and a prolongation $\mathfrak{o}_2$ of $\mathfrak{o}$ in k_2 such that $[X]'$ is rational over k_2 and $\rho_2([X]') = [\mathfrak{X}]$ where ρ_2 is the reduction-mapping with respect to $\mathfrak{o}_2$. By Proposition 9 of Section 1, there exist an extension k_3 of k_1 and a prolongation $\mathfrak{o}_3$ of $\mathfrak{o}_1$ in k_3 such that the residue field k_4 of $\mathfrak{o}_3$ contains k_2. Let $\mathfrak{o}_4$ be a prolongation of $\mathfrak{o}_2$ in k_4. Then we have $\rho_3([X]) = [X]'$, $\rho_4([X]') = [\mathfrak{X}]$ where ρ_3 and ρ_4 are the reduction-mappings with respect to $\mathfrak{o}_3$ and $\mathfrak{o}_4$, respectively. Let $\mathfrak{R}$ be the inverse image of $\mathfrak{o}_4$ by the natural homomorphism of $\mathfrak{o}_3$ onto k_4. Then $\mathfrak{R}$ is a (non-discrete) valuation ring with the quotient field k_3; we may consider the residue field of $\mathfrak{o}_4$ as the residue field of $\mathfrak{R}$. Denote by the $[A]_\lambda$ all the components of $[X]$. We may assume that the $[A]_\lambda$ are all defined over k_3. We fix our attention to one component $[A]_\lambda$ and one of its representative $A_{\lambda\alpha}$. Let S^n be the ambient space for $A_{\lambda\alpha}$, (x) a generic point of $A_{\lambda\alpha}$ over k_3 and the t_{ij} for $1 \leqq i \leqq r$, $1 \leqq j \leqq n$ rn independent variables over k where r is the dimension of $A_{\lambda\alpha}$. Put $y_i = \sum t_{ij}x_j$ for $1 \leqq i \leqq r$. Then we have an irreducible equation $F_{\lambda\alpha}(Y, T) = 0$ for (y_i, t_{ij}) over k_3. For each $[A]_\lambda$ and for each representative $A_{\lambda\alpha}$ of $[A]_\lambda$, we obtain such an equation. We may assume that the polynomials $F_{\lambda\alpha}$ are all primitive polynomials in $\mathfrak{R}[Y, T]$. Denote by c_ν all the coefficients of the $F_{\lambda\alpha}$. Let (γ) be the uniquely determined specialization of (c) over $\mathfrak{R}$. Then obviously we have $(c) \xrightarrow{\mathfrak{o}} (\gamma)$. By Proposition 26, there exists a prolongation $\mathfrak{o}'$ of $\mathfrak{o}$ in $k(c)$ such that $(c) \xrightarrow{\mathfrak{o}'} (\gamma)$. By Proposition 16 of Section 2 and by our definition of the reduction, we can easily prove that $[X]$ is rational over $k(c)$ and $\rho'([X]) = [\mathfrak{X}]$ where ρ' is the reduction-mapping with respect to $\mathfrak{o}'$. This proves the assertion i). The second assertion can be proved similarly.

6. Reduction with respect to infinitely many $\mathfrak{p}$.

Let K be a field with a set of infinitely many discrete valuations $\{\omega_\lambda\}$. We denote by $\mathfrak{o}_\lambda$, $\mathfrak{p}_\lambda$ and κ_λ the valuation ring, the valuation ideal and the residue field of ω_λ, for every λ; denote by σ that set of valuations. In this section, we shall use these notations always in this sense. We assume that the set σ satisfies the following condition:

(I) *Every element of K other than zero is a $\mathfrak{p}_\lambda$-unit for almost all*[11] *the $\mathfrak{p}_\lambda$.*

---

[11] We shall use the term "almost all" in the sense of "all but a finite number of."

In the following we shall use the notations $\sigma_1, \sigma_2, \cdots$ for subsets of σ which contain almost all the $\mathfrak{p}_\lambda$.

Let V be a variety defined over K. We denote by $\bar{V}^{(\lambda)}$ the bunch obtained from V by the reduction with respect to $\mathfrak{p}_\lambda$, for every λ. Similarly, we use the notations $\rho_\lambda(X)$, $[\bar{V}]^{(\lambda)}$ and $\rho_\lambda([X])$ in the sense of the reduction with respect to $\mathfrak{p}_\lambda$.

PROPOSITION 29. *Let U be a variety in $S^n \times S^m$ defined over K with the projection V on S^n. If $[U:V]=1$, then the projection from U to V is regular along every component of $\bar{V}^{(\lambda)}$ for almost all the $\mathfrak{p}_\lambda$.*

Proof. Let (x, y) be a generic point of U over K; then we have expressions $y_j = F_j(x)/F(x)$ where $F(X)$ and $F_j(X)$ are polynomials in $K[X]$. By condition (I), $F(X)$ and the $F_j(X)$ are all contained in $\mathfrak{o}_\lambda[X]$ for almost all the $\mathfrak{p}_\lambda$. Let the t_{ij} $(1 \leqq i \leqq r, 1 \leqq j \leqq n)$ be rn independent variables over $K(x)$ and put $t_i = \sum t_{ij}x_j$ for $1 \leqq i \leqq r$. Let the $\tau_i{}^{(\lambda)}$ and the $\tau_{ij}{}^{(\lambda)}$ $(1 \leqq i \leqq r,\ 1 \leqq j \leqq n)$ be $r(n+1)$ independent variables over κ_λ, for every λ. Put $K' = K(t_i, t_{ij})$, $\mathfrak{o}_\lambda' = [(t_i, t_{ij}) \xrightarrow{\mathfrak{o}_\lambda} (\tau_i{}^{(\lambda)}, \tau_{ij}{}^{(\lambda)})]$; then, by Proposition 3 of Section 1, for every λ, $\mathfrak{o}_\lambda'$ is a discrete valuation ring with the quotient field K'. It is easy to see that the set σ' of all the valuations with the valuation rings $\mathfrak{o}_\lambda'$ also satisfies the condition (I). Since $K'(x)$ is finite algebraic over K', the set σ'' of all the valuations which are prolongations on $K'(x)$ of valuations in σ' also satisfies the condition (I); in particular, $F(x)$ is a unit for almost all valuations in σ''. Let $\mathfrak{V}_\lambda$ be a component of $\bar{V}^{(\lambda)}$ and $\mathfrak{L}_\lambda$ the linear variety $(\sum \tau_{ij}{}^{(\lambda)}X_j - \tau_i{}^{(\lambda)} = 0$ $(1 \leqq i \leqq r))$. Then we have a point $(\xi^{(\lambda)})$ in $\mathfrak{V}_\lambda \cap \mathfrak{L}_\lambda$. The valuation obtained from the specialization $(x) \xrightarrow{\mathfrak{o}_\lambda'} (\xi^{(\lambda)})$ is clearly in σ'', so that $\bar{F}^{(\lambda)}(\xi^{(\lambda)}) = 0$ only for a finite number of the $\mathfrak{p}_\lambda$, where $\bar{F}^{(\lambda)}$ is the class of F modulo $\mathfrak{p}_\lambda$. This proves our proposition.

LEMMA 3. *Let $F(X)$ be an absolutely irreducible polynomial in $K[X]$. Suppose that $F(X)$ is in $\mathfrak{o}_\lambda[X]$ for every λ, and denote by $\bar{F}^{(\lambda)}(X)$ the class of $F(X)$ modulo $\mathfrak{p}_\lambda$. Then $\bar{F}^{(\lambda)}(X)$ is absolutely irreducible for almost all the $\mathfrak{p}_\lambda$.*

Proof. We may assume that the degree of $F(X)$ is greater than 1 and that $\bar{F}^{(\lambda)}(X)$ has the same degree as $F(X)$ for every λ. Let the $M_\alpha{}^{(r)}$ be all the monomials in X whose degrees are less than or equal to r. Let d be the degree of F; and let r and s be a pair of positive integers such that $d = r + s$. Let $G_1(X) = \sum x_\alpha M_\alpha{}^{(r)}$ and $G_2(X) = \sum y_\beta M_\beta{}^{(s)}$ be two polynomials such that (x_α, y_β) is a set of independent variables over K, and put

$$H(X) = \sum z_\gamma M_\gamma{}^{(d)} = G_1(X)G_2(X).$$

We consider (x) and (y) as homogeneous coordinates of generic points of the projective spaces $[L]^{(r)}$ and $[L]^{(s)}$ over K, respectively. Also we consider (z) as homogeneous coordinates of a generic point of some projective variety $[V]$ over K. It is easy to see that (z) has actually a locus over K. By considering the "generic polynomials" over κ_λ, in the same way as above, we have points $(\xi^{(\lambda)})$, $(\eta^{(\lambda)})$ and $(\zeta^{(\lambda)})$ with loci $[\mathfrak{L}]_\lambda{}^{(r)}$, $[\mathfrak{L}]_\lambda{}^{(s)}$ and $[\mathfrak{W}]_\lambda$ over κ_λ, for every λ. It is easy to see that $(\xi^{(\lambda)}) \times (\eta^{(\lambda)}) \times (\zeta^{(\lambda)})$ is a specialization of $(x) \times (y) \times (z)$ over $\mathfrak{o}_\lambda$; so that $[\mathfrak{W}]_\lambda$ is contained in $[\bar{V}]_\lambda$. Now a polynomial $\sum c_\gamma M_\gamma{}^{(d)}$ in $K[X]$ of degree d has a factor of degree r in an extension of K if and only if (c) is a point of $[V]$; and a similar fact holds for a polynomial in $\kappa_\lambda[X]$ and the variety $[\mathfrak{W}]_\lambda$. In fact, the converse part is obvious; so we have only to prove the direct part. If (c) is a point in $[V]$, then (c) is a specialization of (z) over K. This specialization can be extended to a specialization $(a) \times (b) \times (c)$ of $(x) \times (y) \times (z)$ where $(a) \times (b)$ is a point in $[\mathfrak{L}]^{(r)} \times [\mathfrak{L}]^{(s)}$. By the equation $(\sum x_\alpha M_\alpha{}^{(r)})(\sum y_\beta M_\beta{}^{(s)}) = \sum z_\gamma M_\gamma{}^{(d)}$, we can easily verify (if necessary, we consider affine representatives) $(\sum a_\alpha M_\alpha{}^{(r)})(\sum b_\beta M_\beta{}^{(s)}) = h \sum c_\gamma M_\gamma{}^{(d)}$ where h is a quantity. Thus we have proved the above fact. Let $\mathfrak{a}$ be the ideal in $K[Z]$ determined by the set (z). Put $F(X) = \sum c_\gamma M_\gamma{}^{(d)}$; then by our assumption that F is absolutely irreducible, (c) is not a point in $[V]$; so that there exists a homogeneous polynomial $P(Z)$ in $\mathfrak{a}$ such that $P(c) \neq 0$. By our condition (I), the coefficients of $P(Z)$ and $P(c)$ are all $\mathfrak{p}_\lambda$-units for every $\mathfrak{p}_\lambda$ in some $\sigma_1 \subset \sigma$. Put $\bar{F}^{(\lambda)}(X) = \sum \bar{c}_\gamma{}^{(\lambda)} M_\gamma{}^{(d)}$ for every $\mathfrak{p}_\lambda$ in σ_1. If $\bar{F}^{(\lambda)}$ has a factor of degree r, then $(\bar{c}^{(\lambda)})$ is a specialization of $(\zeta^{(\lambda)})$ over κ_λ; so that it is a specialization of (z) over $\mathfrak{o}_\lambda$. Then from $P(z) = 0$ follows $\bar{P}^{(\lambda)}(\bar{c}^{(\lambda)}) = 0$; this is a contradiction if $\mathfrak{p}_\lambda \in \sigma_1$. Now if $\bar{F}^{(\lambda)}(X)$ is reducible, the degree r of one of the factors must satisfy $1 \leqq r \leqq [d/2]$; so our lemma is proved.

PROPOSITION 30. *Let V^r be a variety in S^{r+1} defined over K. Then V is $\mathfrak{p}_\lambda$-simple for almost all the $\mathfrak{p}_\lambda$.*

Proof. By Proposition 2 of [WF] Chapter IV, Section 1, V is defined by an absolutely irreducible equation $F(X) = 0$ with coefficients in K. By our condition (I), the coefficients of $F(X)$ are all $\mathfrak{p}_\lambda$-units for every $\mathfrak{p}_\lambda$ in some $\sigma_1 \subset \sigma$. Denote by $\bar{F}^{(\lambda)}(X)$ the class of $F(X)$ modulo $\mathfrak{p}_\lambda$ for every $\mathfrak{p}_\lambda$ in σ_1. Then by Lemma 3, $\bar{F}^{(\lambda)}$ is absolutely irreducible for every $\mathfrak{p}_\lambda$ in some $\sigma_2 \subset \sigma_1$. By Theorem 21 of Section 4, V is $\mathfrak{p}_\lambda$-simple for every $\mathfrak{p}_\lambda$ in σ_2; this proves our proposition.

THEOREM 26. *Let* $[V]$ *be an abstract variety defined over* K; *let* $[V]_\lambda$ *be a* $\mathfrak{p}_\lambda$*-variety with the underlying abstract variety* $[V]$, *for every* λ. *Suppose that* $[\bar{V}]_\lambda$ *is not empty for almost all the* $\mathfrak{p}_\lambda$. *Then* $[V]_\lambda$ *is* $\mathfrak{p}_\lambda$*-simple for almost all the* $\mathfrak{p}_\lambda$.[12]

Therefore, in this case, for almost all the $\mathfrak{p}_\lambda$, we can obtain the abstract variety $[\mathfrak{V}]_\lambda$ from $[V]_\lambda$ by the reduction with respect to $\mathfrak{p}_\lambda$; and $[\mathfrak{V}]_\lambda$ is defined over κ_λ.

Proof. We shall first prove our theorem in case where $[V]$ is a variety V^r in the affine space S^n. Let (x) be a generic point of V over K. Let the t_{ij} $(0 \leqq i \leqq r, 1 \leqq j \leqq n)$ be $(r+1)n$ independent variables over $K(x)$; and put $K' = K(t)$ and $y_i = \sum t_{ij}x_j$ for $0 \leqq i \leqq r$. Then by Proposition 16 of Section 2, we have $K'(x) = K'(y)$. Let W be the locus of (y) over K' in S^{r+1} and T the birational correspondence between V and W with the corresponding generic points (x) and (y) over K'. Let the $\tau_{ij}^{(\lambda)}$ $(0 \leqq i \leqq r, 1 \leqq j \leqq n)$ be $(r+1)n$ independent variables over κ_λ and put $\mathfrak{o}_\lambda' = [(t) \xrightarrow{\mathfrak{o}_\lambda} (\tau^{(\lambda)})]$. Then by Proposition 3 of Section 1, for every λ, $\mathfrak{o}_\lambda'$ is a discrete valuation ring with the quotient field K'. It is easy to see that the set σ' of all the valuations whose valuation rings are $\mathfrak{o}_\lambda'$ also satisfies (I). By Proposition 30, W is $\mathfrak{p}_\lambda'$-simple for every $\mathfrak{p}_\lambda'$ in some $\sigma_1' \subset \sigma'$. Denote by σ_1 the subset of σ corresponding to the subset σ_1' of σ'. Let $(\xi^{(\lambda)})$ be a generic point of a component $\mathfrak{V}_{\lambda 1}$ of $\bar{V}^{(\lambda)}$ over $\bar{\kappa}_\lambda(\tau^{(\lambda)})$, for every $\mathfrak{p}_\lambda$ in σ_1. Put $\eta_i^{(\lambda)} = \sum \tau_{ij}^{(\lambda)}\xi_j^{(\lambda)}$ for $0 \leqq i \leqq r$; then $(\xi^{(\lambda)}, \eta^{(\lambda)})$ is a specialization of (x, y) over $\mathfrak{o}_\lambda'$; and by Proposition 16 of Section 2, (x) is finite over $(y) \xrightarrow{\mathfrak{o}_\lambda'} (\eta^{(\lambda)})$. Let $\mathfrak{W}_\lambda$ be the variety obtained from W by the reduction with respect to $\mathfrak{p}_\lambda'$, for every $\mathfrak{p}_\lambda'$ in σ_1'. Since $\eta_i^{(\lambda)} = \sum \tau_{ij}^{(\lambda)}\xi_j^{(\lambda)}$, the dimension of $(\eta^{(\lambda)})$ over $\kappa_\lambda(\tau^{(\lambda)})$ is r; so that $(\eta^{(\lambda)})$ is a generic point of $\mathfrak{W}_\lambda$ over $\kappa_\lambda(\tau^{(\lambda)})$. By Theorem 15 of Section 3, $[(y) \xrightarrow{\mathfrak{o}_\lambda'} (\eta^{(\lambda)})]$ is integrally closed, so contains all the x_j since (x) is finite over $(y) \xrightarrow{\mathfrak{o}_\lambda'} (\eta^{(\lambda)})$. This shows that $(\xi^{(\lambda)})$ and $(\eta^{(\lambda)})$ are biregularly corresponding points by T. Then by Theorem 14 of Section 3, we have $\mu(V, \mathfrak{V}_{\lambda 1}) = \mu(W, \mathfrak{W}_\lambda) = 1$. Moreover if $\mathfrak{V}_{\lambda 2}$ is another component of $\bar{V}^{(\lambda)}$ then $\mathfrak{V}_{\lambda 2}$ also biregularly corresponds to $\mathfrak{W}_\lambda$, so it must coincide with $\mathfrak{V}_{\lambda 1}$. Thus our theorem is proved for an affine variety. Now we shall prove our theorem for an abstract variety $[V] = [V_\alpha; F_\alpha; T_{\beta\alpha}]$ on which a $\mathfrak{p}_\lambda$-variety $[V]_\lambda = [V_\alpha; F_\alpha; \mathfrak{F}_\alpha^{(\lambda)}; T_{\beta\alpha}]$ is defined for every λ. As our theorem is proved for affine varieties, the varieties V_α and $T_{\beta\alpha}$ are all $\mathfrak{p}_\lambda$-simple for every $\mathfrak{p}_\lambda$ in some $\sigma_1 \subset \sigma$. Denote by $\mathfrak{V}_\alpha^{(\lambda)}$ the variety obtained from V_α by the reduction with respect to $\mathfrak{p}_\lambda$, for every α and every $\mathfrak{p}_\lambda$ in σ_1. By Proposition 29, for every α and β, $T_{\beta\alpha}$ is regular along $\mathfrak{V}_\alpha^{(\lambda)}$ and $\mathfrak{V}_\beta^{(\lambda)}$ for every $\mathfrak{p}_\lambda$ in some $\sigma_2 \subset \sigma_1$. Then by Theorem 13 of Section 3, there exists a component $\mathfrak{T}_{\beta\alpha}^{(\lambda)}$ of $\bar{T}_{\beta\alpha}^{(\lambda)}$ having the projection $\mathfrak{V}_\alpha^{(\lambda)}$. Since $T_{\beta\alpha}$ is $\mathfrak{p}_\lambda$-simple if $\mathfrak{p}_\lambda \varepsilon \sigma_1$, $\mathfrak{T}_{\beta\alpha}^{(\lambda)}$ is the only component of $\bar{T}_{\beta\alpha}^{(\lambda)}$ if $\mathfrak{p}_\lambda \varepsilon \sigma_2$; so that, by the same reason, the projection of $\mathfrak{T}_{\beta\alpha}^{(\lambda)}$ on $\bar{V}_\beta^{(\lambda)}$ is equal to $\mathfrak{V}_\beta^{(\lambda)}$. This shows that $\mathfrak{V}_\alpha^{(\lambda)}$ and $\mathfrak{V}_\beta^{(\lambda)}$ are biregularly corresponding varieties by $T_{\beta\alpha}$ for every $\mathfrak{p}_\lambda$ in σ_2. From this follows that, for every $\mathfrak{p}_\lambda$ in σ_2, $[\bar{V}]_\lambda$ has only one component $[\mathfrak{V}]_\lambda$. Furthermore, by our definition we have $\mu([V]_\lambda, [\mathfrak{V}]_\lambda) = \mu(V_\alpha, \mathfrak{V}_\alpha^{(\lambda)}) = 1$, if $[\mathfrak{V}]_\lambda$ has a representative in V_α. This proves our theorem.

---

[12] This theorem may be considered as a generalization of a theorem of Bertini for pencils. (See O. Zariski, "Pencils on an algebraic variety and a new proof of a theorem of Bertini," *Transactions of the American Mathematical Society*, vol. 50 (1941), pp. 48-70.)

UNIVERSITY OF TOKYO.

REFERENCES.

[1] M. Deuring, "Reduktion algebraischer Funktionenkörper nach Primdivisoren des Konstantenkörpers," *Mathematische Zeitschrift*, vol. 47 (1942), pp. 643-654.

[2] ———, "Die Struktur der elliptischen Funktionenkörper und die Klassenkörper der imaginären quadratischen Zahlkörper," *Mathematische Annalen*, vol. 124 (1952), pp. 393-426.

[3] W. Krull, "Dimensionstheorie in Stellenringen," *Journal für die Reine und Angewandte Mathematik*, vol. 179 (1938), pp. 204-226.

[4] T. Matsusaka, "Specialization of cycles on a projective model," *Memoirs of the College of Science, University of Kyoto*, vol. 26 (1950), pp. 167-173.

[5] D. G. Northcott, "Specialization over a local domain," *Proceedings of the London Mathematical Society*, (III), vol. 1 (1951), pp. 129-137.

[6] P. Samuel, "La notion de multiplicité en algèbre et en géométrie algébrique," *Journal de Mathématique Pures et Appliquées*, (IX), vol. 30 (1951), pp. 159-274.

[7] A. Weil, *Foundations of algebraic geometry*, New York, 1946.

[8] ———, "Arithmetic on algebraic varieties," *Annals of Mathematics*, vol. 53 (1951), pp. 412-444.

COHERENCE INVARIANT MAPPINGS ON KRONECKER PRODUCTS.*

By Henry G. Jacob, Jr.

---

1. Introduction. The main objective in this paper is a generalization of a theorem on matrices proved by L. K. Hua [2]. By means of the concept of the rank of a matrix Hua has formulated the notion of coherence between pairs of matrices and defined the class of one-to-one coherence invariant mappings on the group $n \times m$ matrices. His theorem gives a complete characterization of this class for the case of matrices with entries in a field. The extension of Hua's results is suggested by the relationship between linear transformations of a finite dimensional space and matrices. Since in the finite dimensional setting the group of linear transformations is a special case of the broader notion of Kronecker products it appears natural to attempt to phrase a generalization in terms of Kronecker products of arbitrary dimensional spaces over division rings. This is accomplished by stating new definitions of rank and coherence for Kronecker products which reduce in the special case to the older forms. The generalized results are given in Theorem 5.1 of Section 5, and the last section is devoted to applications of this theorem.

2. The Kronecker product. As mentioned in the introduction, the system which serves as a foundation for this investigation is that of the Kronecker product of two vector spaces. We shall assume that $\mathfrak{X}$ and $\mathfrak{Y}$ are arbitrary left and right vector spaces respectively over the division ring Δ.

Definition 2.1. *$\mathfrak{X} \otimes \mathfrak{Y}$ is called the Kronecker product of the left vector space $\mathfrak{X}$ and the right vector space $\mathfrak{Y}$ if and only if,*

(i) *$\mathfrak{X} \otimes \mathfrak{Y}$ is a commutative group (composition denoted by $+$).*

(ii) *there exists a mapping $(x, y) \rightarrow x \otimes y$ of the product set $\mathfrak{X} \times \mathfrak{Y}$ into $\mathfrak{X} \otimes \mathfrak{Y}$ which, for all $x, x_1, x_2 \in \mathfrak{X}$, $y, y_1, y_2 \in \mathfrak{Y}$, $\alpha \in \Delta$ satisfies*

---

* Received September 30, 1953; revised November 6, 1954.

(a) $(x_1 + x_2) \otimes y = x_1 \otimes y + x_2 \otimes y$.

(b) $x \otimes (y_1 + y_2) = x \otimes y_1 + x \otimes y_2$.

(c) $x\alpha \otimes y = x \otimes \alpha y$.

(iii) *Every element of* $\mathfrak{X} \otimes \mathfrak{Y}$ *is of the form* $\sum x_i \otimes y_i$, $x_i \in \mathfrak{X}$, $y_i \in \mathfrak{Y}$.

(iv) (a) *If* $e_1, e_2, \cdots, e_m$ *is a finite set of linearly independent vectors in* $\mathfrak{X}$, *then the relation* $\sum^m_1 e_j \otimes y_j = 0$ *holds for* $y_j \in \mathfrak{Y}$ *only if all* y_j*'s* $= 0$.

(b) *If* $f_1, f_2, \cdots, f_n$ *is any finite set of linearly independent vectors in* $\mathfrak{Y}$, *then the relation* $\sum^n_1 x_j \otimes f_j = 0$ *holds for* $x_j \in \mathfrak{X}$ *only if all the* x_j*'s* $= 0$.

Properties (iv) (a) and (b) are called the independence conditions and it can be shown that any two groups which possess properties (i)-(iv) are isomorphic. A realization of this group is obtained by letting $\mathfrak{X}'$ be a left vector space dual to $\mathfrak{X}$ relative a bilinear form $(x', x), x' \in \mathfrak{X}', x \in \mathfrak{X}$ [4]. For $x \in \mathfrak{X}$ and $y \in \mathfrak{Y}$ we define $x \otimes y$ to be the linear transformation $x' \to (x', x)y$ of $\mathfrak{X}'$ into $\mathfrak{Y}$. Then the set $\mathfrak{B} = [\sum x_i \otimes y_i \mid x_i \in \mathfrak{X}, y_i \in \mathfrak{Y}]$ is the Kronecker product of $\mathfrak{X}$ and $\mathfrak{Y}$.

3. Rank and coherence.

DEFINITION 3. 1. *Let* $\mathfrak{X}$ *and* $\mathfrak{Y}$ *be right and left vector spaces respectively over the division ring* Δ. *Let* $\mathfrak{X} \otimes \mathfrak{Y}$ *denote the Kronecker product of these spaces. An element* $T = \sum x_i \otimes y_i$ *is of rank* n *if* n *is the minimum number of terms of the form* $x \otimes y$ *in which* T *can be expressed.*

DEFINITION 3. 2. *Two elements* T *and* R *of* $\mathfrak{X} \otimes \mathfrak{Y}$ *are said to be coherent if* $T - R$ *has rank one.*

In the above realization of the Kronecker product as linear transformations, it is evident that rank as defined coincides with the more familiar notion for transformations. Now if we denote the rank of an element T by $r(T)$ we have as an obvious fact, that if T and R are elements of $\mathfrak{X} \otimes \mathfrak{Y}$, $r(T + R) \leqq r(T) + r(R)$.

DEFINITION 3. 3. *If* $(x_1, x_2, \cdots, x_n)$ *is a set of vectors in* $\mathfrak{X}$, *the space generated by these will be denoted by* $[x_1, x_2, \cdots, x_n]$. *A similar notation will be used for subspaces generated by elements of* $\mathfrak{Y}$.

LEMMA 3. 1. *An element* $T \in \mathfrak{X} \otimes \mathfrak{Y}$ *is of rank* n *if and only if* $T = \sum^n_1 x_i \otimes y_i$ *where the sets* $(x_1, x_2, \cdots, x_n)$ *and* $(y_1, y_2, \cdots, y_n)$ *are* *linearly independent. Furthermore, if* T *is of rank* n *and* $T = \sum^{n}_{1} x_i \otimes y_i = \sum^{n}_{1} x_i^* \otimes y_i^*$, *then*

$$[x_1, x_2, \cdots, x_n] = [x_1^*, x_2^*, \cdots, x_n^*]$$

and

$$[y_1, y_2, \cdots, y_n] = [y_1^*, y_2^*, \cdots, y_n].$$

Proof. The necessity of the first statement is obvious. To prove the sufficiency assume that $T = \sum^{n}_{1} x_i \otimes y_i$, where $(x_1, x_2, \cdots, x_n)$ and $(y_1, y_2, \cdots, y_n)$ each form linearly independents sets. Suppose that $r(T) < n$ and $T = -\sum^{n+k}_{n+1} x_j \otimes y_j$, $k < n$, where $(x_{n+1}, \cdots, x_{n+k})$ and $(y_{n+1}, \cdots, y_{n+k})$ are linearly independent sets. Then $\sum^{n+k}_{1} x_i \otimes y_i = 0$. From the set $(y_1, y_2, \cdots, y_{n+k})$ select the maximal linearly independent subset which comprises the first n elements. We have then $(y_1, y_2, \cdots, y_n, y_{n+1}, \cdots, y_{n+s})$, $s \leqq k$ and $y_t = \sum^{n+s}_{1} \gamma_j^t y_j$, $t = n+s+1, n+s+2, \cdots, n+k$, hence

$$\sum_{i=1}^{n+s} x_i \otimes y_i + \sum_{i=n+s+1}^{n+k} x_i \otimes (\sum_{j=1}^{n+s} \gamma_j^i y_j) = 0 \quad \text{or} \quad \sum_{j=1}^{n+s} [x_j + (\sum_{i=n+s+1}^{n+k} x_i \gamma_j^i)] \otimes y_j = 0.$$

By the assumed independence of $(y_1, y_2, \cdots, y_n, y_{n+1}, \cdots, y_{n+s})$, condition (iv)(a), mentioned earlier, implies that $x_j + (\sum^{n+k}_{n+s+1} x_i \gamma_j^i) = 0$ for $j = 1, 2, \cdots, n+s$. Since $k < n$, the independence of the set $(x_1, x_2, \cdots, x_n)$ is contradicted.

Assume next that T is of rank n and $T = \sum^{n}_{1} x_i \otimes y_i = \sum^{n}_{1} x_i^* \otimes y_i^*$. For simplicity write $x_k^* = -x_{n+k}$ and $y_k^* = y_{n+k}$, for $k = 1, 2, \cdots, n$. Then $T = \sum^{n}_{1} x_i \otimes y_i = -\sum^{2n}_{n+1} x_j \otimes y_j$ which implies $\sum^{2n}_{1} x_i \otimes y_i = 0$. Suppose that $[y_1, \cdots, y_n] \neq [y_1^*, y_2^*, \cdots, y_n^*]$ and $(y_1, y_2, \cdots, y_n, y_{n+1}, \cdots, y_{n+s})$ $0 < s \leqq n$ is a maximal linearly independent set selected from the y's. Following the same argument used in the preceding part of the proof we may show that the independence of the set $(x_1, x_2, \cdots, x_n)$ is contradicted.

The last lemma states that the sets $(x_1, x_2, \cdots, x_n)$ and $(y_1, y_2, \cdots, y_n)$ which appear in a minimal representation of an element T, where $r(T) = n$, will serve as basis for the spaces generated by any other two sets used in a minimal representation of T. The subspace $[x_1, x_2, \cdots, x_n]$ will be called *the right space associated with* T and the subspace $[y_1, y_2, \cdots, y_n]$ will be called *the left space associated with* T. A useful corollary of Lemma 3. 2 is the following

COROLLARY 3. 1. $r(X - Y) = n$, *for* X *and* $Y \varepsilon \mathfrak{X} \otimes \mathfrak{Y}$, *if and only if there exists a chain of* $n - 1$ *elements* $X_1, X_2, \cdots, X_{n-1}$ *in* $\mathfrak{X} \otimes \mathfrak{Y}$ *such that*

$1 = r(X_1 - X) = r(Y - X_{n-1}) = r(X_i - X_{i-1})$, $1 < i < n$, *and there is no shorter chain with this property.*

We shall now begin a study of one-to-one mappings of one Kronecker product onto another which are coherence invariant, i. e. *a mapping is coherence invariant if whenever two elements are coherent their images are coherent.* We shall also require that the inverses of these mappings be coherence invariant. To establish the existence of such a mapping we shall describe two distinct types. Let $\mathfrak{X}_i$ and $\mathfrak{Y}_i$, $i = 1, 2$, be right and left vector spaces respectively over the division ring Δ_i. Suppose that P and Q are semi-linear transformations of $\mathfrak{X}_1$ into $\mathfrak{X}_2$ and $\mathfrak{Y}_1$ into $\mathfrak{Y}_2$ respectively with the same associated isomorphism of Δ_1 onto Δ_2. It is then true that the mapping $\sum x_i \otimes y_i \rightarrow \sum x_i P \otimes y_i Q$ of $\mathfrak{X}_1 \otimes \mathfrak{Y}_1$ into $\mathfrak{X}_2 \otimes \mathfrak{Y}_2$, which we shall denote by $P \otimes Q$, is single-valued.* In the case where P and Q are one-to-one and onto $P \otimes Q$ is also one-to-one and coherence invariant, furthermore, the inverse $P^{-1} \otimes Q^{-1}$ possesses these properties. A second type of coherence invariant mapping different from the first is obtained in the following manner. Suppose that Δ_1' is an anti-isomorphic image of Δ_1, then the spaces $\mathfrak{X}_1$ and $\mathfrak{Y}_1$ may be regarded as left and right vector spaces respectively over Δ_1', and, as such, have associated with them the Kronecker product $\mathfrak{Y}_1 \square \mathfrak{X}_1$. It can be easily shown that the mapping $\sum x_i \otimes y_i \rightarrow \sum y_i \square x_i$ *is* a natural isomorphism of $\mathfrak{X}_1 \otimes \mathfrak{Y}_1$ onto $\mathfrak{Y}_1 \square \mathfrak{X}_1$, which we shall denote by δ. Now if Q and P are one-to-one semi-linear transformations of $\mathfrak{Y}_1$ onto $\mathfrak{X}_2$ and $\mathfrak{X}_1$ onto $\mathfrak{Y}_2$ with the same isomorphisms of Δ_1' onto Δ_2, the resultant mapping $\sum x_i \otimes y_i \rightarrow \sum y_i \square x_i \rightarrow \sum y_i Q \otimes x_i P$ is both one-to-one and coherence invariant of $\mathfrak{X}_1 \otimes \mathfrak{Y}_1$ onto $\mathfrak{X}_2 \otimes \mathfrak{Y}_2$, and the same is true of the inverse.

We observe that the mappings of the two classes already described always carry the zero element of $\mathfrak{X}_1 \otimes \mathfrak{Y}_1$ onto the zero of $\mathfrak{X}_2 \otimes \mathfrak{Y}_2$. Therefore, it should be remarked that a slightly more general coherence invariant mapping is obtained by the addition of a constant element in the image space. For example, consider the mapping $T \rightarrow T[P \otimes Q] + R$, where R is a fixed element of $\mathfrak{X}_2 \otimes \mathfrak{Y}_2$. Clearly this mapping has essentially the same properties as $P \otimes Q$.

At this point the question which naturally arises is whether or not the two classes described above comprise the totality of one-to-one coherence invariant mappings which have coherence invariant inverses and take zero onto zero. The answer, which is given in Theorem 5. 1, is in the affirmative when the dimensionalities of the vector spaces involved are greater than or equal to two and the division rings have more than two elements.

---

* This can be shown by an argument similar to that for the direct product of linear transformations [4].

4. The fundamental theorem of affine geometry.

Essential to the proof of the last statement of Section 3 is a well known theorem in affine geometry which we shall call the "fundamental theorem of affine geometry." The theorem will be stated for right vector spaces only, although similar results obtain for left vector spaces. The finite dimensional case over a field with three or more elements has been treated by Hua [2].

DEFINITION 4. 1. *If* $\mathfrak{X}$ *is a right vector space over a division ring* Δ *and* $x_1, x_2 \,\varepsilon\, \mathfrak{X}$, *the set of points* $\mathfrak{L} = [x_1\alpha + x_2(1 - \alpha) \mid \alpha \,\varepsilon\, \Delta]$ *is called the line through* x_1 *and* x_2.

THEOREM 4. 1. *Let* $\mathfrak{X}_1$ *and* $\mathfrak{X}_2$ *be right vector spaces over* Δ_1 *and* Δ_2 *respectively, with dimensionality greater than or equal to two. If* Δ_1 *contains more than two elements, then every one-to-one mapping* P *of* $\mathfrak{X}_1$ *onto* $\mathfrak{X}_2$ *which carries lines onto lines and zero onto zero is a semi-linear transformation.*

A comparison of Theorem 4. 1 with the "fundamental theorem of projective geometry" [1] will reveal that the minimum dimension in this case is one lower, however, an additional condition must be imposed on the division ring. The following example will show that this restriction cannot be removed. Let $\mathfrak{X}$ be a three dimensional vector space over Φ, the field of two elements. Denote by (x_1, x_2, x_3) a basis for $\mathfrak{X}$. It is clear that $\mathfrak{X}$ contains a total of eight elements which enumerated are $0, x_1, x_2, x_3, x_1 + x_2, x_1 + x_3, x_2 + x_3$, and $x_1 + x_2 + x_3$. Consider the mapping P on $\mathfrak{X}$ which maps 0 onto $0, x_1 + x_2$ onto $x_1 + x_3$, $x_1 + x_3$ onto $x_1 + x_2$ and acts as the identity on the remaining elements. Since any line in $\mathfrak{X}$ contains just two elements (in particular a line determined by two points contains only these points) it is quickly seen that P is one-to-one on $\mathfrak{X}$ and carries lines onto lines. However, P is obviously not additive.

5. The main theorem.

The notion of coherent sets, which is the fundamental tool in the analysis of coherence invariant mappings, will be introduced now.

DEFINITION 5. 1. (i) *A set* $\mathfrak{S}$ *contained in the Kronecker product* $\mathfrak{X} \otimes \mathfrak{Y}$ *is said to be a coherent set if every pair of elements of* $\mathfrak{S}$ *is coherent.*

(ii) *a maximal coherent set is a coherent set which is not properly contained in any other coherent set. Maximal coherent sets may frequently be referred to as simply maximal sets.*

(iii) *a maximal set containing the* 0 *element* (*hence, in addition, only elements of rank one*) *will be called a maximal set of rank one.*

DEFINITION 5.2. *We shall denote by* $\mathfrak{J}$ *the group of mappings of* $\mathfrak{X} \otimes \mathfrak{Y}$ *of the form* $T \rightarrow T_R = T + R$, R *a fixed element of* $\mathfrak{X} \otimes \mathfrak{Y}$. *Two sets of* $\mathfrak{X} \otimes \mathfrak{Y}$ *are said to be equivalent under* $\mathfrak{J}$ *if one is the image of the other under some element of* $\mathfrak{J}$.

These mappings are analogous to translations in a vector space and are certainly one-to-one and coherence invariant. The same is true of the inverse, hence they leave invariant maximal sets and intersections of maximal sets.

NOTATION. (i) *If* $x \varepsilon \mathfrak{X}$ *and* $y \varepsilon \mathfrak{Y}$ *we shall denote by* $x \otimes \mathfrak{Y}$ *and* $\mathfrak{X} \otimes y$ *the sets* $[x \otimes y \mid y \varepsilon \mathfrak{Y}]$ *and* $[x \otimes y \mid x \varepsilon \mathfrak{X}]$ *respectively.*

(ii) *Since* $T \rightarrow T^\delta$ *denotes the natural isomorphism of* $\mathfrak{X}_1 \otimes \mathfrak{Y}_1$ *onto* $\mathfrak{Y}_1 \square \mathfrak{X}_1$, *if* $T = \sum x_i \otimes y_i$ *then* $T^\delta = \sum y_i \square x_i$. *The resultant mapping* $T \rightarrow T^\delta \rightarrow \sum y_i Q \otimes x_i P$, *described earlier, will be denoted by* $T \rightarrow (T^\delta)^\eta$.

THEOREM 5.1. *Let* $T \rightarrow T^\sigma$ *be a one-to-one coherence invariant mapping of* $\mathfrak{X}_1 \otimes \mathfrak{Y}_1$ *onto* $\mathfrak{X}_2 \otimes \mathfrak{Y}_2$ *which carries* 0 *onto* 0 *and possesses a coherence invariant inverse. Then either*

(1) $T^\sigma = \sum x_i P \otimes y_i Q$, *where* P *and* Q *are semi-linear transformations of* $\mathfrak{X}_1$ *onto* $\mathfrak{X}_2$ *and* $\mathfrak{Y}_1$ *onto* $\mathfrak{Y}_2$ *respectively with the same associated isomorphisms of* Δ_1 *onto* Δ_2, *or*

(2) $T^\sigma = (T^\delta)^\eta = \sum y_i Q \otimes x_i P$, *where* P *and* Q *are semi-linear transformations of* $\mathfrak{X}_1$ *onto* $\mathfrak{Y}_2$ *and* $\mathfrak{Y}_1$ *onto* $\mathfrak{X}_2$ *with the same associated isomorphisms of* Δ'_1 *onto* Δ_2, *regarding* $\mathfrak{X}_1$ *and* $\mathfrak{Y}_1$ *as left and right spaces over* Δ'_1.

We shall first establish the theorem for elements of rank one by examining the effect of σ on maximal sets of rank one. Two lemmas are required for this part of the proof.

LEMMA 5.1. *In the Kronecker product* $\mathfrak{X} \otimes \mathfrak{Y}$ *maximal sets of rank one are either of the form* $x \otimes \mathfrak{Y}$ *or* $\mathfrak{X} \otimes y$ *where* x *and* y *are non-zero elements of* $\mathfrak{X}$ *and* $\mathfrak{Y}$ *respectively.*

Proof. It follows from Lemma 3.1 that if Y is coherent to both 0 and

$X = x \otimes y \neq 0$, an element of rank one, then either $Y = x \otimes v$ or $Y = u \otimes y$. Thus if $\mathfrak{S}$ is a maximal set of rank one containing X, every element of $\mathfrak{S}$ is of one of these two types. Since $\mathfrak{S}$ is maximal, and dimension of $\mathfrak{X}$ and $\mathfrak{Y} \geqq 2$, it contains either a $u \otimes y$ where u and x are linearly independent or $x \otimes v$ where v and y are linearly independent. If both belong, then $\mathfrak{S}$ is no longer a coherent set.

LEMMA 5. 2. *There are exactly two maximal sets containing a given pair of coherent elements. If the coherent elements* $X = x_1 \otimes y$ *and* $X_2 = x_2 \otimes y$ *are contained in the maximal sets* $\mathfrak{S}_1$ *and* $\mathfrak{S}_2$, *then*

$$\mathfrak{S}_1 \cap \mathfrak{S}_2 = [x_1\lambda + x_2(1-\lambda) \otimes y \mid \lambda \varepsilon \Delta].$$

Proof. Let X_1 and X_2 be any pair of coherent elements. Choose from $\mathfrak{T}$ the mapping $T \rightarrow T - X_2$; then $X_1 \rightarrow x \otimes y$, an element of rank one, and $X_2 \rightarrow 0$. Now $x \otimes \mathfrak{Y}$ and $\mathfrak{X} \otimes y$ are the only maximal sets containing 0 and $x \otimes y$, hence the sets $\mathfrak{S}_1$ and $\mathfrak{S}_2$, equivalent to $x \otimes Y$ and $X \otimes y$ under this mapping, are the only maximal sets containing X_1 and X_2. To prove the second statement suppose $X_1 = x_1 \otimes y$ and $X_2 = x_2 \otimes y$ and apply the same mapping again. Then $\mathfrak{S}_1 \cap \mathfrak{S}_2$ is equivalent to the set $[(x_1 - x_2)\lambda \otimes y \mid \lambda \varepsilon \Delta]$.

Proof of the Theorem. From the assumptions about the mapping $T \rightarrow T^\sigma$, it is immediate that maximal sets of rank one and intersections of maximal sets are left invariant. As a consequence of this, if x_1 and w_1 are linearly independent elements of $\mathfrak{X}_1$, and $(x_1 \otimes \mathfrak{Y}_1)^\sigma$ and $(w_1 \otimes \mathfrak{Y}_1)^\sigma$ are images of $x_1 \otimes \mathfrak{Y}_1$ and $w_1 \otimes \mathfrak{Y}_1$ respectively, then either $(x_1 \otimes \mathfrak{Y}_1)^\sigma = x_2 \otimes \mathfrak{Y}_2$ and $(w_1 \otimes \mathfrak{Y}_1)^\sigma = w_2 \otimes \mathfrak{Y}_2$ where x_2, w_2 are linearly independent, or $(x_1 \otimes \mathfrak{Y}_1)^\sigma = \mathfrak{X}_2 \otimes y_2$ and $(x_1 \otimes \mathfrak{Y}_1) = \mathfrak{X}_2 \otimes z_2$ where y_2, z_2 are linearly independent in $\mathfrak{Y}_2$. Furthermore, if $(x_1 \otimes \mathfrak{Y}_1)^\sigma = x_2 \otimes \mathfrak{Y}_2$, then $(\mathfrak{X}_1 \otimes y_1)^\sigma = \mathfrak{X}_2 \otimes y_2$ must occur. Now it will be more convenient to restrict the remaining discussion to the case in which $(x_1 \otimes \mathfrak{Y}_1)^\sigma = x_2 \otimes \mathfrak{Y}_2$, however, it should be understood that the following arguments apply equally well to the other possible situation.

Since σ is one-to-one we may write $(x_1 \otimes y_1)^\sigma = x_2 \otimes y_1^\beta$, where $y_1 \rightarrow y_1^\beta$ is a one-to-one mapping of $\mathfrak{Y}_1$ onto $\mathfrak{Y}_2$. In fact, β possesses the additional property of carrying lines onto lines, which is seen by considering the line $[\lambda y_1 + (1-\lambda)z_1 \mid \lambda \varepsilon \Delta_1]$ and recognizing, by Lemma 5. 2, the set $[x_1 \otimes \lambda y_1 + (1-\lambda)z_1 \mid \lambda \varepsilon \Delta_1]$ as the intersection of the maximal sets containing the coherent pair $x_1 \otimes y_1$ and $x_1 \otimes z_1$. The latter set, it follows, must map onto $[x_2 \otimes \gamma y_1^\beta + (1-\gamma)z_1^\beta \mid \gamma \varepsilon \Delta_2]$ which is the intersection of the maximal sets containing the images $x_2 \otimes y_1^\beta$ and $x_2 \otimes z_1^\beta$. Thus by Theorem 4. 1 we may assert that for each $x_1 \varepsilon \mathfrak{X}_1$ the mapping $y_1 \rightarrow y_1^\beta$ is a one-to-one semi-linear transformation of $\mathfrak{Y}_1$ onto $\mathfrak{Y}_2$. Furthermore, from the coherence invariance of σ and the linearity of these induced mappings it can be seen that any two differ only by a scalar multiple. Similarly for each $y_1 \varepsilon \mathfrak{Y}_1$ there corresponds a one-to-one semi-linear transformation of $\mathfrak{X}_1$ onto $\mathfrak{X}_2$, and these also differ by scalars. Consequently, we may select a single mapping from each set, denote them by P_1 and Q_1, and write $(x_1 \otimes y_1)^\sigma = x_1P_1 \otimes \gamma(x_1, y_1)y_1Q_1$ for all elements of rank one, where $\gamma(x_1, y_1)$ appears to be dependent on x_1 and y_1. Actually $\gamma(x_1, y_1)$ is independent, and to show this we need only observe that the mappings $y_1 \rightarrow \gamma(x_1, y_1)y_1Q$, for x_1 fixed, and $x_1 \rightarrow x_1P_1\gamma(x_1, y_1)$, for y_1 fixed, are also semi-linear. Thus if P is the semi-linear transformation $x_1 \rightarrow x_1P_1\gamma(x_1, y_1)$ and $Q = Q_1$, we have $(x_1 \otimes y_1)^\sigma = x_1P \otimes y_1Q$. Clearly P and Q will have the same isomorphisms.

With the theorem established for elements of rank one we proceed with the final step. To simplfy this step we introduce the composite mapping $T \rightarrow T^\sigma[P^{-1} \otimes Q^{-1}]$ and denote it by π. π is one-to-one and coherence invariant on $\mathfrak{X}_1 \otimes \mathfrak{Y}_1$, leaving the elements of rank one pointwise fixed. We shall prove that it is the identity on all of $\mathfrak{X}_1 \otimes \mathfrak{Y}_1$. Suppose that T is of rank n. By Lemma 3.1, if $T = \sum^n_1 x_i \otimes y_i$, the sets $(x_1, x_2, \cdots, x_u)$ and $(y_1, y_2, \cdots, y_u)$ are each linearly independent. It is immediate by Corollary 3.1 that $r(T) = r(T^\pi)$, hence we may write $T^\pi = \sum^n_1 x_j^* \otimes y_j^*$, where the x^*'s and y^*'s also form linearly independent sets. We observe next that $T - x_i \otimes y_i$, $1 \leqq i \leqq n$, is of rank $n-1$, therefore by Corollary 3.1 $T^\pi - (x_i \otimes y_i)^\pi = T^\pi - x_i \otimes y_i$ is of rank $n-1$. Thus $T^\pi = x_i \otimes y_i + \sum^{n-1}_1 x_{ik} \otimes y_{ik}$ for $1 \leqq i \leqq n$, and we have at once that $x_i \varepsilon [x_1^*, \cdots, x_n^*]$ and $y_i \varepsilon [y_1^*, y_2^*, \cdots, y_n^*]$. We may now write $T^\pi = \sum^n_1 x_j \otimes v_j$, where $(v_1, \cdots, v_n)$ is a linearly independent set in $\mathfrak{Y}_1$.

Consider the element $T - x_1 \otimes (y_1 + \lambda_2y_2 + \cdots + \lambda_ny_n)$ which is of rank $n-1$ for all λ_j, $j = 2, 3, \cdots, n$. Then $T^\pi - x_1 \otimes (y_1 + \lambda_2y_2 + \cdots + \lambda_ny_n)$ is of rank $n-1$, which implies $v_1 - y_1 = \sum^n_2 \xi_iv_i + \lambda_2y_2 + \cdots + \lambda_ny_n$ for all λ_j. Let $\lambda_j = 0$, $j = 2, \cdots, n$, then $v_1 - y_1 = \sum^n_2 \xi_{1i}^{(1)}v_i$. By letting $\lambda_j = 0, j \neq k$, the last two equations give $y_k = \sum^n_2 \xi_{1i}^{(k)}v_i$, $k \neq 1$. Applying a similar argument to $T - x_j \otimes (\lambda_1y_1 + \lambda_2y_2 + \cdots + y_j + \cdots + \lambda_ny_n)$ we obtain

$$v_j - y_j = \sum_{i=1}^{n}{}' \xi_{ji}^{(j)}v_i \quad \text{and} \quad y_k = \sum_{i=1}^{n}{}' \xi_{ji}^{(k)}v_i, \qquad k \neq j,$$

where the primes refer to the exclusion of the summation index $i = j$. From these relations we have, for example, if $j = 2, 3, \cdots, n$,

$$v_1 = y_1 + \sum_{i=2}^{n} \xi_{1i}^{(1)} v_i = \sum_{i=1}^{n}{}' \xi_{ji}^{(1)} v_i + \sum_{i=2}^{n} \xi_{1i}^{(1)} v_i.$$

Since the v_i's are linearly independent, by letting j run from 2 to n we see $\xi_{1i}^{(1)} = 0$, $i = 2, 3, \cdots, n$, hence $v_1 = y_1$. By the same argument it can be shown that $v_i = y_i$, $i = 1, 2, \cdots, n$. Consequently, $T^\pi = T$ for all $T \varepsilon \mathfrak{X}_1 \otimes \mathfrak{Y}_1$ and $T^\sigma = T(P \otimes Q) = \sum^n{}_1 x_i P \otimes y_i Q$.

6. Applications. We turn our attention next to a situation in which we are given the dual pairs of vectors spaces $(\mathfrak{X}', \mathfrak{X})$ and $(\mathfrak{Y}', \mathfrak{Y})$ over the division ring Δ, with associated nondegenerate bilinear forms (x', x) $x' \varepsilon \mathfrak{X}'$, $x \varepsilon \mathfrak{X}$ and (y', y), $y' \varepsilon \mathfrak{Y}'$, $y \varepsilon \mathfrak{Y}$. We shall consider the additive group $\mathfrak{L}(\mathfrak{X}', \mathfrak{Y}')$ of linear transformations of $\mathfrak{X}'$ into $\mathfrak{Y}'$ which are continuous in the $\mathfrak{X}$-topology of $\mathfrak{X}'$ and $\mathfrak{Y}$-topology of $\mathfrak{Y}'$. It can be shown that these are just the linear transformations which possess adjoint mappings of $\mathfrak{Y}$ into $\mathfrak{X}$[4]. Moreover, the set $\mathfrak{F}(\mathfrak{X}', \mathfrak{Y}')$ of linear transformations of the form

$$x' \to (x', x_1) y_1' + (x', x_2) y_2' + \cdots + (x', x_n) y_n', \; x' \varepsilon \mathfrak{X}', x_1 \varepsilon \mathfrak{X}, y_i' \varepsilon \mathfrak{Y}', \quad 1 \leqq i \leqq n,$$

is a subset of $\mathfrak{L}(\mathfrak{X}', \mathfrak{Y}')$ which coincides with the set of all continuous linear transformations of finite rank. A transformation of this type will be denoted by $\sum x_i \otimes y_i'$. We have already seen that $\mathfrak{F}(\mathfrak{X}', \mathfrak{Y}')$ is the Kronecker product $\mathfrak{X} \otimes \mathfrak{Y}'$.

It should be mentioned that $\mathfrak{L}(\mathfrak{X}', \mathfrak{Y}')$ is a topological group in the finite topology, and $\mathfrak{F}(\mathfrak{X}', \mathfrak{Y}')$ is a dense subgroup of the full group of linear transformations of $\mathfrak{X}'$ into $\mathfrak{Y}'$. In addition $\mathfrak{L}(\mathfrak{X}', \mathfrak{Y}')$ may be considered as a left or right module over $\mathfrak{L}(\mathfrak{X}', \mathfrak{X}')$ or $\mathfrak{L}(\mathfrak{Y}', \mathfrak{Y}')$ respectively, where $\mathfrak{L}(\mathfrak{X}', \mathfrak{X}')$ is the ring of linear transformations in $\mathfrak{X}'$ with adjoints, and $\mathfrak{L}(\mathfrak{Y}', \mathfrak{Y}')$ is similarly defined for $\mathfrak{Y}'$. The subring $\mathfrak{F}(\mathfrak{X}', \mathfrak{X}')$ is a simple two sided ideal of $\mathfrak{L}(\mathfrak{X}', \mathfrak{X}')$ comprised of the union of minimal right ideals. Each minimal right ideal is generated by a idempotent transformation of rank one, hence the minimal ideals are composed of transformations of rank one and every transformation of rank one belongs to a minimal right ideal. Since $\mathfrak{L}(\mathfrak{X}', \mathfrak{Y}')$ is a left module over $\mathfrak{L}(\mathfrak{X}', \mathfrak{X}')$ we have the multiplication rule

$$(x_1 \times x_1')(x_2 \otimes y_2') = x_1(x_1', x_2) \otimes y_2',$$

where $x_1 \times x_1' \varepsilon \mathfrak{X} \times \mathfrak{X}' = \mathfrak{F}(\mathfrak{X}', \mathfrak{X}')$ and $x_2 \otimes y_2' \varepsilon \mathfrak{X} \otimes \mathfrak{Y}' = \mathfrak{F}(\mathfrak{X}', \mathfrak{Y}')$.

Now consider the left module $\mathfrak{A}$ over $\mathfrak{R}$ such that $\mathfrak{F}(\mathfrak{X}', \mathfrak{Y}') \subseteqq \mathfrak{A} \subseteqq \mathfrak{L}(\mathfrak{X}', \mathfrak{Y}')$, $\mathfrak{F}(\mathfrak{X}', \mathfrak{X}') \subseteqq \mathfrak{R} \subseteqq \mathfrak{L}(\mathfrak{X}', \mathfrak{X}')$, and $\mathfrak{R}\mathfrak{A} \subseteqq \mathfrak{A}$. By the density property we have

LEMMA 6.1. *If R, $S \varepsilon \mathfrak{R}$ and $RA = SA$ for all $A \varepsilon \mathfrak{A}$, then $R = S$. If $A, B \varepsilon \mathfrak{A}$ and $RA = RB$ for all $R \varepsilon \mathfrak{R}$, then $A = B$.*

Our purpose in this section is to investigate a one-to-one mapping σ of the module $\mathfrak{A}_1$ over $\mathfrak{R}_1$ onto the module $\mathfrak{A}_2$ over $\mathfrak{R}_2$, ($\mathfrak{A}_i$ and $\mathfrak{R}_i$, $i = 1, 2$, as described above) which satisfies the condition that $(R_1A_1)^\sigma = R_1{}^\gamma A_1{}^\sigma$ where $R_1 \rightarrow R_1{}^\gamma$ is a one-to-one mapping of $\mathfrak{R}_1$ onto $\mathfrak{R}_2$. In the process of characterizing this type of mapping we shall show that the associated mapping $R_1 \rightarrow R_1{}^\gamma$ is multiplicative, hence additive by a theorem of C. E. Rickart [5]. Moreover, an extension of Rickart's methods will be used to prove the additivity of σ. It will be assumed that the vector spaces and division rings involved satisfy the conditions imposed in Section 5.

LEMMA 6.2. *Let $A_1 \rightarrow A_1{}^\sigma$ be a one-to-one mapping of $\mathfrak{A}_1$ onto $\mathfrak{A}_2$ such that $(R_1A_1)^\sigma = R_1{}^\gamma A_1{}^\sigma$ where $R_1 \rightarrow R_1{}^\gamma$ is a one-to-one mapping of $\mathfrak{R}_1$ onto $\mathfrak{R}_2$. Then the mapping σ is additive and γ is an isomorphism.*

Proof. Since $(R_1A_1)^\sigma = R_1{}^\gamma A_1{}^\sigma$ it follows that $(R_1S_1)^\gamma A_1{}^\sigma = R_1{}^\gamma S_1{}^\gamma A_1{}^\sigma$ for all $A_1 \varepsilon \mathfrak{A}_1$, hence by Lemma 6.1 $(R_1S_1)^\gamma = R_1{}^\gamma S_1{}^\gamma$. Thus γ is multiplicative which implies γ is additive and, in fact, an isomorphism. Now since any element of rank one is contained in a minimal right ideal and minimal right ideals are mapped onto minimal right ideals under isomorphisms, we see that the additivity of γ implies the coherence invariance of γ. Furthermore, γ^{-1} is coherence invariant by the same argument, consequently $\mathfrak{F}(\mathfrak{X}_1', \mathfrak{Y}_1')$ is mapped onto $\mathfrak{F}(\mathfrak{X}_2', \mathfrak{Y}_2')$.

Next suppose A_1 and $B_1 \varepsilon \mathfrak{F}(\mathfrak{X}_1', \mathfrak{Y}_1')$ and have disjoint right spaces associated with minimal representations; i.e., $A_1 = \sum^{n}{}_1 x_i \otimes y_i'$ and $B_1 = \sum^{m}{}_1 u_i \otimes v_i'$, where $(x_1, x_2, \cdots, x_n, u_1, \cdots, u_m)$ form a linearly independent set. Let $x_{n+k} = u_k$, $k = 1, 2, \cdots, m$, then by the assumption that $(\mathfrak{X}_1', \mathfrak{X}_1)$ comprise a dual pair of vector spaces there exist $x_1', x_2', \cdots, x_{n+m}'$, $x_i' \varepsilon \mathfrak{X}_1'$ such that $(x_i', x_j) = \delta_{ij}$. If $R_1 = \sum^{n}{}_1 x_i \times x_i'$ and $S_1 = \sum^{m+n}{}_{n+1} x_j \times x_j'$, it is apparent that $R_1A_1 = A_1$, $R_1B_1 = 0$, $S_1B_1 = B_1$ and $S_1A_1 = 0$. Thus by the additivity of γ and the equation $(R_1 + S_1)(A_1 + B_1) = A_1 + B_1$, we have $(A_1 + B_1)^\sigma = A_1{}^\sigma + B_1{}^\sigma$. Furthermore, an induction argument may be used to show that $A_1{}^\sigma = \sum^{n}{}_1 (x_i \otimes y_i)^\sigma$, whenever the x_i's are linearly independent. Consider next the case in which $A_1 = x_1 \otimes y'$ and $B_1 = x_1 \otimes v'$. Since $\mathfrak{X}_1$ has dimension greater than or equal to two there exists $x_1{}^* \varepsilon \mathfrak{X}_1$ such that x_1 and $x_1{}^*$ are linearly independent. If $C_1 = x_1{}^* \otimes y$, then $A_1 + B_1 = x_1 \otimes (y' + v')$ and $A_1 + C_1 = (x_1 + x_1{}^*) \otimes y$, thus from the preceding remarks it follows that

$$((A_1 + B_1) + C_1)^\sigma = (A_1 + B_1)^\sigma + C_1^\sigma$$

$$((A_1 + C_1) + B_1)^\sigma = (A_1 + C_1)^\sigma + B_1^\sigma = A_1^\sigma + C_1^\sigma + B_1^\sigma.$$

Therefore $(A_1 + B_1)^\sigma = A_1^\sigma + B_1^\sigma$.

We are now prepared to deal with the case in which A_1 and B_1 are arbitrary elements of $\mathfrak{F}(\mathfrak{X}_1', \mathfrak{Y}_1')$. Assume that the set $(x_1, x_2, \cdots, x_{n+1} \cdots x_{n+m})$ is not linearly independent and select from it the maximal linearly independent set $(x_1, x_2, \cdots, x_n, x_{n+1}, \cdots, x_{n+t})$, where $0 \leqq t < m$. Then we may write $B_1 = \sum^n{}_1\, x_i \otimes w_i' + \sum^{n+t}{}_{n+1}\, x_j \otimes w_j'$ and

$$A_1 + B_1 = \sum^n{}_1\, x_i \otimes (y_i' + w_i') + \sum^{n+t}{}_{n+1}\, x_j \otimes w_j'.$$

From these equations it is easily shown that σ is additive on $\mathfrak{F}(\mathfrak{X}_1', \mathfrak{Y}_1')$.

Finally, to show σ is additive on all of $\mathfrak{A}_1$, it is sufficient, in view of Lemma 6.1, to observe that for A_1 and B_1 in $\mathfrak{A}_1$, $R_1^\gamma(A_1^\sigma + B_1^\sigma) = R_1^\gamma(A_1 + B_1)^\sigma$ for all $R_1 \varepsilon \mathfrak{F}(\mathfrak{X}_1', \mathfrak{X}_1')$.

THEOREM 6.1. *Let* $A_1 \rightarrow A_1^\sigma$ *be as defined in Lemma* 6.1. *Then* $A_1^\sigma = T_1^* A_1 Q_1$ *and* $R_1^\gamma = T_1^* R_1 T_1^{*-1}$, *where* T_1^* *and* Q_1 *are one-to-one semi-linear transformations of* $\mathfrak{X}_2'$ *onto* $\mathfrak{X}_1'$ *and* $\mathfrak{Y}_1'$ onto $\mathfrak{Y}_2'$ *respectively.*

Proof. In the proof of Lemma 6.2 we showed γ is coherence invariant. We must also show the same is true of σ. Suppose $L_1 = x_1 \otimes y_1'$. There exists $F_1 = x_1 \times x_1'$ such that $(x_1, x_1') = 1$, hence $F_1 L_1 = L_1$ and $(L_1)^\sigma = F_1^\gamma L_1^\sigma$. But F_1 is of rank one, consequently $F_1^\gamma L_1^\sigma$ is of rank one. Since σ is also additive it must be coherent invariant. σ^{-1} is coherence invariant for the same reasons.

We may now conclude by Theorem 5.1 that for all

$$F_1 = \sum x_i \times x_i' \varepsilon \mathfrak{F}(\mathfrak{X}_1', \mathfrak{X}_1')$$

either (1) $F_1^\gamma = \sum x_i T_1 \times x_i' T_1'$ or (2) $F_1^\gamma = \sum x_i' T_1' \times x_i T_1$, and for all $L_1 \varepsilon \mathfrak{F}(\mathfrak{X}_1', \mathfrak{Y}_1')$ either (3) $L_1^\sigma = \sum x_j P \otimes y_j' Q$ or (4) $L_1^\sigma = \sum y_j' Q \otimes x_j P$. To start with we shall assume that forms (1) and (3) hold simultaneously. Later we show that this is the only possible combination.

Suppose that $F_1 = x_1 \times x_1'$ and $L_1 = x_1 \otimes y_1'$. Then

$$F_1 L_1 = x_1 (x_1', x_1) \otimes y_1', \quad \text{hence} \quad (F_1 L_1)^\sigma = (x_1 (x_1', x_1)) P \otimes y_1' Q.$$

By assumption

$$(F_1 L_1)^\sigma = F_1^\gamma L_1^\sigma = (x_1 T_1 \times x_1' T_1')(x_1 P \otimes y_1' Q)$$
$$= x_1 T_1 (x_1' T_1', x_1 P) \otimes y_1' Q,$$

and thus $x_1P(x_1', x_1)^\tau = x_1T_1(x_1'T_1', x_1P)$, where τ is the isomorphism associated with P. Since P and T are one-to-one and differ for each $x_1 \varepsilon \mathfrak{X}_1$ by a scalar multiple, we have that $x_1P = x_1T_1\alpha$, α fixed in Δ_2. Moreover, if λ is the isomorphism associated with T_1, then $\gamma^\lambda = \alpha\gamma^\tau\alpha^{-1}$, $\gamma \varepsilon \Delta_1$. Thus $\alpha(x_1', x_1)^\tau = (x_1'T_1', x_1T_1)\alpha$, from which follows $(x_1', x_1)^\lambda = (x_1'T_1', x_1T_1)$. Since T_1 and T_1' have inverses, the last relation states

$$(x_2', x_1T_1)^{\lambda^{-1}} = (x_2'T_1'^{-1}, x_2).$$

In other words, T_1 has as adjoint $T_1^* = T_1'^{-1}$, consequently if $y_1Q_1 = \alpha y_1Q$, then $(F_1L_1)^\sigma = T_1^*(F_1L_1)Q_1$, $F_1^\gamma = T_1^*F_1T_1^{*-1}$, and $L_1^\sigma = T_1^*L_1Q_1$ for all L_1 and F_1. Again Lemma 6. 1 can be used to prove that σ takes this form for all $A_1 \varepsilon \mathfrak{A}_1$. Moreover, since it is evident that γ is, in fact, a special case of σ when we identify $\mathfrak{A}_1$ with $\mathfrak{R}_1$, it is also clear that γ has the same form on all of R_1.

Finally we may dispose of the other combinations by considering only the combination (1) and (4) and remarking that a similar argument can be used for the others. Suppose then that F_1 and L_1 are defined as before and $(F_1L_1)^\sigma = y_1'Q \otimes (x_1(x_1', x_1))P$ and $F_1^\gamma L_1^\sigma = x_1T_1(x_1'T_1', y_1'Q) \otimes x_1P$. Therefore, $y_1'Q(x_1', x_1)^\tau = x_1T_1(x_1'T_1', y_1'Q)$ for all $y_1' \varepsilon \mathfrak{Y}_1'$, $x_1 \varepsilon \mathfrak{X}_1$, $x_1' \varepsilon \mathfrak{X}_1'$. Clearly if y_1 is fixed and x_1 allowed to vary, this situation is impossible. This completes the proof.

By means of Theorem 6. 1 we may readily obtain the isomorphism theorem for primitive rings with minimal ideals [3]. This result can be stated as follows.

THEOREM 6. 2. *Let* $(\mathfrak{X}_1', \mathfrak{X}_1)$ *and* $(\mathfrak{X}_2', \mathfrak{X}_2)$ *be dual pairs of vector spaces over* Δ_1 *and* Δ_2 *respectively. Let* $\mathfrak{A}_i$, $i = 1, 2$, *be primitice rings with minimal ideals such that* $\mathfrak{F}(\mathfrak{X}_i', \mathfrak{X}_i) \subseteqq \mathfrak{A}_i \subseteqq \mathfrak{L}(\mathfrak{X}_i', \mathfrak{X}_i')$. *If* $A_1 \rightarrow A_1^\sigma$ *is an isomorphism of* $\mathfrak{A}_1$ *onto* $\mathfrak{A}_2$, *then* $A_1^\sigma = T_1^*A_1T_1^{*-1}$, *where* T_1^* *is a one-to-one semi-linear transformation of* $\mathfrak{X}_2'$ *onto* $\mathfrak{X}_1'$.

LOUISIANA STATE UNIVERSITY.

BIBLIOGRAPHY.

[1] R. Baer, *Linear algebra and projective geometry*, New York, 1952.

[2] L. K. Hua, "A theorem on matrices and its application to Grassmann space," *Science Reports of the National Tsing Hua University*, Ser. A, vol. 5 (1948), pp. 150-181.

[3] N. Jacobson, "On the theory of primitive rings," *Annals of Mathematics*, vol. 48 (1947), pp. 8-21.

[4] ———, *Lectures in abstract alegbra*, vol. I, New York (1951), vol. II, New York (1953).

[5] C. E. Rickart, "One-to-one mappings of rings and lattices," *Bulletin of the American Mathematical Society*, vol. 54 (1948), pp. 758-764.

A CHARACTERIZATION OF A CLASS OF RINGS.*

By R. L. San Soucie.

---

1. Introduction. Skornyakov [4] has shown that *a right alternative division ring of characteristic not two is alternative,* while the author has proved [2] that *a right alternative division ring of characteristic two is alternative if and only if it satisfies the identity*

$$w(xy \cdot x) = (wx \cdot y)x. \tag{1.1}$$

Moreover, R. H. Bruck has actually constructed examples of right alternative division rings which are not alternative. This construction is detailed in an appendix to [2], but may be summarized as follows.

Let F be any field of characteristic two for which there exists an automorphism α of order 2 and an element f_0 in F such that $f_0\alpha = f_0$ and f_0 not a square in F. Let θ be the additive endomorphism of F defined by $f\theta = f\alpha + f_0 f$, all f in F. Then if R is the set of all couples (f, g), with f, g in F, if equality and addition in R are defined componentwise and multiplication is defined by

$$(f, g)(h, k) = (fh + g \cdot k\theta, fk + gh), \tag{1.2}$$

R is a right alternative division ring of characteristic two which is not alternative. A field F which suffices is $F = B(t)$, where B is any field of characteristic two and t is transcendental over B. Then, for any $f(t)$ in $B(t)$, α defined by $f(t)\alpha = f(t+1)$ and $f_0 = t^2 + t$ can be used to construct the required endomorphism θ. (The elaborate choice of θ is explained in [2], and does not concern us here.)

In view of the above, we make the following definition. If B is a not alternative, right alternative division ring of characteristic two, if B is two-dimensional over some field F (so that B can be written as the set of all couples (f, g) with f, g in F) and if multiplication in B is given by (1.2), where θ is some additive endomorphism of F, not necessarily, however, the particular kind mentioned above, then we shall call B a *Bruck ring.* Since the rings whose construction was outlined above fall into this class, the class is not empty. Moreover, Bruck rings have the following surprising properties:

---

* Received November 8, 1954.

1. The identity $(w, x, yz) + (w, y, xz) = (w, x, z)y + (w, y, z)x$ holds in any Bruck ring B, although it can be obtained by linearization from (1. 1), which cannot hold in B.
2. The identity $(wx, y, z) = w(x, y, z) + (w, y, z)x + (w, x, (y, z))$ is satisfied by any choice of w, x, y, z in any B, and this identity played an important role in proving that other right alternative rings were alternative. (See [1], [2], [3], and [4].)
3. The left nucleus K, of any B, commutes with B, and squares and commutators are all in K.

In view of these properties, we concern ourselves in this paper with the following problem: what characterizes the Bruck rings in the class of all not alternative, right alternative division rings of characteristic two? The answer turns out to be simple. Property 2 and a part of property 3 (above) together provide necessary and sufficient conditions for a not alternative right alternative division ring to be a Bruck ring.

All of our results appear to be new, for (to our knowledge) there is no literature concerning right alternative rings that does not either prove or assume (1. 1). But some of our methods are similar to those used in [2]. We intentionally omit the verification of properties 1, 2, and 3 for Bruck rings. Nothing is involved except tedious computation using (1. 2) and this does not seem worth recording here.

2. Preliminary results.

Throughout the paper, R will always denote a right alternative ring of characteristic two.

LEMMA 1. *The following identities hold in R:*

$$(wx, y, z) + (w, xy, z) + (w, x, yz) = w(x, y, z) + (w, x, y)z \tag{2. 1}$$

$$(xy, z) = x(y, z) + (x, z)y + (z, x, y) \tag{2. 2}$$

$$((x, y), z) + ((x, z), y) + ((y, z), x) = 0. \tag{2. 3}$$

These identities are well-known. (See [5], page 125.)

Definition 1. R will be said to have *property A* if R is not alternative and if

$$(wx, y, z) = w(x, y, z) + (w, y, z)x + (w, x, (y, z)) \tag{2. 4}$$

holds identically in R.

We now derive some consequences of the definition.

Lemma 2. *Let R have property A. Then the following identities hold in R:*

$$(w,x,yz)+(w,y,xz)=(w,x,z)y+(w,y,z)x \tag{2.5}$$

$$(w,x,(y,z))+(w,y,(x,z))+(w,z,(x,y))=0 \tag{2.6}$$

$$(w,x^2,y)=(w,x,y)x+(w,x,yx) \tag{2.7}$$

$$(w,y,x)\cdot yx+(w,y,x)x\cdot y=(w,y,yx^2)+(w,y,x^2)y. \tag{2.8}$$

Proof. (2.5) is obtained by direct substitution of (2.4) into (2.1) and collecting terms. (2.6) then follows from (2.5) by expansion of the left member of (2.6). (2.7) then comes from (2.5) via the substitution $x\rightarrow y$, $y\rightarrow x$, $z\rightarrow x$. Now $(w,y,yx^2)=(w,y,x)\cdot yx+(w,yx,x)y$ using (2.5) and also $(w,yx,x)y=(w,y,x^2)y+(w,y,x)x\cdot y$ using (2.7). This proves (2.8) and the lemma.

The following definition is standard.

Definition 2. Let

$$K=[k\,\varepsilon\,R\mid(k,R,R)=0],\text{ let } M=[m\,\varepsilon\,R\mid(R,R,m)=0],$$

and let $C=[c\,\varepsilon\,M\mid(c,R)=0]$. Then we call K the *left nucleus* of R, M the *right nucleus* of R, and C the *centre* of R. Observe that C is in K by (2.2).

Lemma 3. *Let R have property A and have no proper divisors of zero. Then $M=C$.*

Proof. Clearly we need only prove $M\subset C$. Select $x\,\varepsilon\,R$, $x\not\varepsilon\,M$. Then, for any $m\,\varepsilon\,M$, $(R,x,m)=0$ and $(x,m)\,\varepsilon\,M$, using (2.4). But also $(R,x,xm)=0$ so that $(x,xm)\,\varepsilon\,M$. However, $(xm,x)=x(m,x)$ and the latter implies $x\,\varepsilon\,M$, a contradiction, unless $(m,x)=0$. Hence $0=(xm,m')=x(m,m')$ and thus $(M,R)=0$. This proves the lemma.

Definition 3. If R is a ring with property A and if R is such that all commutators are in the left nucleus, then we shall say that R also has the *commutator property*. For convenience, we shall write that R has *property ACP*.

Lemma 4. *Let R have property ACP and no proper divisors of zero. Then* (i) $(K,K)=0$, (ii) $(x^2,R)=0$, (iii) $x^2\,\varepsilon\,K$, (iv) $((x,y),R)=0$, (v) $(R,x^2,(y,z))=0$, (vi) $(R,x^2,y^2)\subset K$.

Proof. Select $x\,\varepsilon\,R$, $x\not\varepsilon\,K$, and $k,k'\,\varepsilon\,K$. Then

$$(xk, k') = x(k, k') + (x, k')k$$

and thus $x(k, k') \varepsilon K$. But the latter implies that $x \varepsilon K$, a contradiction, unless $(k, k') = 0$. This proves (i). Since

$$(2.9) \qquad (xy, x) = x(y, x) + (x, x, y),$$

we have

$$((x, x, y), x, y) = (x(y, x), x, y) = (x, x, y)(x, y)$$

and also

$$((x, x, y), x, y) = ((y, x)x, x, y) = (x, y)(x, x, y)$$

so that

$$(2.10) \qquad ((x, y), (x, x, y)) = 0.$$

We conclude from (2.10) and (2.9) that $0 = (x(x, y), (x, y)) = (x, (x, y))(x, y)$ and this implies that $(x, (x, y)) = 0$. Hence $(x^2, y) = x(x, y) + (x, y)x = 0$ and $(x^2, R) = 0$. (ii) and (2.2) prove (iii), while (iv) follows from (ii) by replacing x by $x + y$. Replacing x by x^2 in (2.6) then yields (v), and (vi) comes from the observation that $(x^2y^2, u) = (u, x^2, y^2)$.

Henceforth we shall assume that R is a division ring with property *ACP*. This added assumption enables us to prove a lemma which is crucial to our theory. It is perhaps worth noting that the identity of the lemma has heretofore been proved using (1.1).[1]

LEMMA 5. *The following identity holds in R:*

$$(2.11) \qquad ((w, x, y), x, y) = (w, x, y)(x, y).$$

Proof. We begin by observing that

$$(w, y, yx^2) = (w, y, x^2y) = (w, x^2, y^2) + (w, x^2, y)y.$$

Hence (2.8) becomes

$$(2.12) \qquad (w, y, x) \cdot yx + (w, y, x)x \cdot y = (w, x^2, y^2).$$

Then

$$\begin{aligned} ((w, x, y), x, y) &= (w, x, y) \cdot xy + (w, x, y)x \cdot y \\ &= (w, x, y) \cdot xy + (w, x, y) \cdot yx + (w, x^2, y^2), \end{aligned}$$

using (2.12), so that

$$(2.13) \qquad ((w, x, y), x, y) = (w, x, y)(x, y) + (w, x^2, y^2).$$

---

[1] But we should point out that previous proofs did not need the division ring hypothesis.

We now complete the proof by showing that $(w, x^2, y^2) = 0$, and begin by noting that (2.9) implies

$$((x, x, y), x, y) = (x, x, y)(x, y), \tag{2.14}$$

while (2.13) yields

$$((x, x, y), x, y) = (x, x, y)(x, y) + (x, x^2, y^2). \tag{2.15}$$

Therefore (2.14) and (2.15) together yield

$$(x, x^2, y^2) = 0. \tag{2.16}$$

The substitution $x \rightarrow x + y$, $y \rightarrow z$ in (2.16) yields the linearized identity

$$(y, x^2, z^2) = (x, y^2, z^2) \tag{2.17}$$

and setting $y = k$ in (2.17) gives

$$(x, k^2, z^2) = 0. \tag{2.18}$$

Now consider $q = (w, x^2, y^2)$. If either x or y is in K, $q = 0$ by (2.18). Hence we can assume without loss of generality that $x \not\in K$ and then we define z by $xz = w$, for the particular x and w chosen in the associator q. This gives

$$q = (xz, x^2, y^2) = x(z, x^2, y^2). \tag{2.19}$$

But $q \in K$ by Lemma 4(vi), and (z, x^2, y^2) is in K also. Hence (2.19) implies that x is in K, a contradiction, unless $(z, x^2, y^2) = 0$. But then $q = 0$, (2.13) becomes (2.11) and the lemma is proved.

3. The characterization.

Since we are assuming that R is not alternative, we known that R contains at least two elements a and b such that $d = (a, b) \neq 0$. If $(R, a, b) = 0$, then $d' = (a, ab) = ad \neq 0$. If also $(R, a, ab) = 0$, then, since d, d' are in M, a is in M and $d = 0$ by Lemma 3, a contradiction. Hence a, b may be chosen so that $(a, b) = d \neq 0$ and $(R, a, b) \neq 0$. With a, b so chosen and fixed once for all, define the mapping π (of R into R) by

$$x\pi = (x, a, b), \qquad \text{all } x \text{ in } R. \tag{3.1}$$

Then we may write (2.4) as

$$(wx)\pi = w \cdot x\pi + w\pi \cdot x + (w, x, d). \tag{3.2}$$

THEOREM 1. *Let*

$$N = [n \varepsilon R \mid n\pi = (n, a, b) = 0], \quad S = [s \varepsilon R \mid s\pi = (s, a, b) = sd].$$

Then $R = N + S$, $S \cap N = 0$, $NS = S$, $N \neq 0 \neq S$, and $S^2 = N$.

Proof. (i) Since $d \,\varepsilon\, K \subset N$, $N \neq 0$. Also, S contains $R\pi$ by Lemma 5 so that $S \neq 0$. Now, for any $x \,\varepsilon\, R$, define q by $qd = x\pi$ and observe that $(qd)\pi = q\pi \cdot d$ while $x\pi \cdot \pi = x\pi \cdot d = qd \cdot d$. Hence $q\pi = qd$ and $q \,\varepsilon\, S$. Set $y = x - q$ and obtain $y\pi = x\pi - q\pi = 0$ so $y \,\varepsilon\, N$, and $R = N + S$. Clearly $S \cap N = 0$.

(ii) Note that $(ns)\pi = n \cdot sd + (n, s, d) = ns \cdot d$ so $NS \subset S$. If x is the solution of $nx = s$, $n \neq 0$, then $x = n' + s'$, $n(n' + s') = s$ and $nn' \,\varepsilon\, S$. Thus $nn' \cdot d = (n, n', d)$ so $n' = 0$, $x = s'$ and $NS = S$.

(iii) Finally, $(ss')\pi = sd \cdot s' + s \cdot s'd + (s, s', d) = s \cdot ds' + s \cdot s'd = 0$ by Lemma 4(iv) so that $S^2 \subset N$. The trick which worked in (ii) may be used again to get $S^2 = N$.

LEMMA 6. *With definitions as in Theorem* 1, (i) $ns \cdot d = s \cdot dn$, (ii) $(dn, s) = 0$, (iii)[2] $(N, R, d) = 0$.

Proof. Since $(n, s)\pi = 0$, we have $ns \cdot d = (sn)\pi = sd \cdot n + (s, d, n)$ and this gives (i). Now $dn \cdot s \,\varepsilon\, S$ and $s \cdot dn \,\varepsilon\, S$ by (i) so that $(s, dn) \,\varepsilon\, S$. But $(s, dn) \,\varepsilon\, K \subset N$ so $(s, dn) = 0$. For (iii) we observe that $(n, x, d) = (nx)\pi + n \cdot x\pi$ so (n, x, d) in S. But also $(n, x, d) = (xd, n) + (x, n)d$ and thus $(n, x, d) \,\varepsilon\, K \subset N$. Hence $(n, x, d) = 0$ and this is (iii).

THEOREM 2. *Let N and S be defined as in Theorem* 1. *Then* (i) $N^2 = N$, (ii) $(N, S) = 0$, (iii) $SN = S$, (iv) $N = K$, (v) $(K, R) = 0$ *and K is a field.*

Proof. Lemma 6(iii) implies that $(nn')\pi = 0$ and hence $N^2 \subset N$. The usual maneuver gives $N^2 = N$. Now we observe that any $n' \,\varepsilon\, N$ can be written $n' = dn$ for some n in N. Then, for any s in S, $(n', s) = (dn, s) = 0$ by Lemma 6(ii). Thus $(N, S) = 0$ and $SN = S$.

Using (ii), $0 = (nn', s) = (s, n, n')$ so $(S, N, N) = 0$. Therefore $0 = (sn, n', n'') = s(n, n', n'')$ and thus $(N, N, N) = 0$, $(R, N, N) = 0$. Now, as in the proof of Lemma 3, $(R, n, n') = 0$ implies $(n, n') \,\varepsilon\, M$. But also $(R, n, nn') = 0$ so $(n, nn') \,\varepsilon\, M$. Hence, if $(n, n') \neq 0$, $n \,\varepsilon\, M$, $(n, n') = 0$ by Lemma 3. This contradiction proves $(N, N) = 0$, $(N, R) = 0$. Thus $N \subset K$ by (2.2) and this proves $N = K$. We need only remark that K is a commutative associative division ring to complete the proof.

We collect our results in

---

[2] We are indebted to the referee for the remark, and the proof, that $(N, R, d) = 0$, thus simplifying our original version of this part of the lemma.

THEOREM 3. *Let R be a not alternative right alternative division ring of characteristic two with the properties that*

$$(wx, y, z) = w(x, y, z) + (w, y, z)x + (w, x, (y, z))$$

and $(x, y) \in K$, the left nucleus of R, for all w, x, y, z of R. Then $R = K + Ks$, K a field, commutative with all of R, and $s \in R$, $s \notin K$. Moreover, multiplication in R is given by

$$(k_1 + k_2 s)(k_3 + k_4 s) = k_1 k_3 + k_2 \cdot k_4 \theta + (k_1 k_4 + k_2 k_3)s,$$

where θ is an additive endomorphism of K.

Proof. The proof is immediate in view of the foregoing and the observation that $k_4\theta = s \cdot sk_4$ is an additive endomorphism of K.

4. Remarks. It would be interesting to know if all not alternative right alternative division rings have property *ACP* of definition 3. And it would be of interest to determine whether there exist any finite Bruck rings. It is clear, of course, that none can be constructed by the method outlined in the introduction, for no finite field can have characteristic two and also a non-square element.

THE UNIVERSITY OF OREGON.

REFERENCES.

[1] E. Kleinfeld, "Right alternative rings," *Proceedings of the American Mathematical Society*, vol. 4 (1953), pp. 939-944.

[2] R. L. San Soucie, "Right alternative division rings of characteristic two," *ibid.*, to appear.

[3] ———, "Right alternative rings of characteristic two," submitted to *Proceedings of the American Mathematical Society.*

[4] L. A. Skornyakov, "Right alternative fields," *Izvestia Akademiia Nauk SSSR. Seriia Matematichesaia,* vol. 15 (1951), pp. 177-184.

[5] Max Zorn, "Theorie der alternativen Ringe," *Abhandlungen Mathematik Seminar Hamburgischen Universitat,* vol. 8 (1930), pp. 123-147.

CALCU
CENTR
LIBRAR
UNIVERS

ON THE GALOIS COHOMOLOGY OF UNRAMIFIED EXTENSIONS OF FUNCTION FIELDS IN ONE VARIABLE.*

By Y. KAWADA and J. TATE.

I. The Abstract Case.

Let k be an algebraic function field in one variable over an algebraicly closed constant field k_0. Let K/k be a finite normal unramified extension with Galois group $G = G(K/k)$. Our aim is to investigate the cohomology groups of the Galois group G with coefficients in various multiplicative groups associated with K. Some of the coefficient groups which we consider, and our notations for them, are the following:

$D(K) =$ the group of divisors of K;

$P(K) =$ the group of principal divisors of K;

$E(K) = D(K)/P(K) =$ the group of divisor classes of K;

$U =$ the group of all roots of unity in k_0;

$E^*(K) = \mathrm{Hom}(E(K), U) =$ the group of all homomorphisms of $E(K)$ into U.

We omit K in these notations when considering a fixed field K.

A rough outline of our program is the following. We first show $H^r(G, E) \approx H^{r+2}(G, U)$ for all r, by using three exact coefficient group sequences linking E to U. By duality it follows then that $H^r(G, E^*) \approx H^{r-2}(G, Z)$, at least in case the characteristic p does not divide the order of G. This result is exactly parallel to the main theorem of class field theory which states $H^r(G, C) \approx H^{r-2}(G, Z)$ if G is a group of automorphisms of a number field with idèle class group C. Therefore, casting the group E^* of divisor class characters in the role of the idèle class group, we can produce a pseudo class field theory for the extensions under consideration. We develop the parallel with ordinary class field thory in some detail, and discuss the relation of our pseudo class field theory to the simpler Kummer theory. In the second half of the paper, we consider the classical case in which the ground field is the

---

[*] Received August 4, 1954.

field of complex numbers. Then the divisor class group E has a natural locally compact topology, and it turns out that its Pontrjagin character group $\hat{E}$ can be used instead of E^*. Viewing 1-cycles on the Riemann surface as characters of the group of divisor classes of degree 0, we can give an explicit topologico-analytic description of the fundamental generator of $H^2(G, \hat{E})$, and can show that the reciprocity law of our pseudo class field theory is the mapping which attaches to a 1-cycle on the ground surface the corresponding covering transformation of an abelian unramified covering surface.

The cohomology groups which we will be using are those associated with the operation of a *finite* group G on a G-module A. They are denoted by $H^r(G, A)$ and are defined, and in general non-trivial, for all integers r, positive and negative. The negative dimensional cohomology groups are isomorphic to the ordinary homology groups of G in A. The group $H^0(G, A)$ may be identified with A^G/NA, the factor group of elements $a \varepsilon A$ which are invariant under the operations of G, modulo elements of the form $a = Nb = \prod_{\sigma \varepsilon G} b^\sigma$ for $b \varepsilon A$. The basic properties of these groups are developed in Chapter XII of [1]. The main fact, which justifies the introduction of negative dimensional groups, is that an exact sequence of coefficient modules of the type $0 \to A' \to A \to A'' \to 0$ leads to an exact cohomology sequence

$$\cdots \to H^r(G, A) \to H^r(G, A'') \to H^{r+1}(G, A') \to H^{r+1}(G, A) \to \cdots$$

which extends all the way from $r = -\infty$ to $r = +\infty$. Furthermore, cup products are defined in all dimensions, positive and negative: if A and B are G-paired into C, there is induced a pairing of $H^r(G, A)$ and $H^s(G, B)$ into $H^{r+s}(G, C)$ for *all* r, s. This pairing has all the usual properties of a cup product. Finally, we mention that $H^r(G, A) = 0$ for all r if A is G-regular; that is, if $A = \sum_{\sigma \varepsilon G} B^\sigma$ is the direct sum of translations of a subgroup $B \subset A$. After these preliminary remarks, we now return to the field theoretic situation described in the first paragraph.

In order to determine the cohomology groups $H^r(G, E(K))$ we consider three natural exact sequences:

$$\{1\} \to P \to D \to E \to \{1\}, \tag{1}$$

$$\{1\} \to k_0^* \to K^* \to P \to \{1\}, \tag{2}$$

$$\{1\} \to U \to k_0^* \to k_0^*/U \to \{1\}. \tag{3}$$

In each of these three sequences, one of the modules has trivial cohomology groups in all dimensions. In the first sequence we have $H^r(G, D) = 0$ for all r, because D is G-regular. The G-regularity of D follows from our assumptions that K/k is unramified and that the constant field k_0 is algebraicly closed. Namely, if for each prime divisor $\mathfrak{p}$ of k we select a fixed prime $\mathfrak{P}$ of K dividing $\mathfrak{p}$, then the other primes of K dividing $\mathfrak{p}$ are those of the form $\mathfrak{P}^\sigma$, for $\sigma \varepsilon G$, and these are distinct. Consequently, if D_1 is the subgroup of D generated by the selected $\mathfrak{P}$'s, then D is the direct product of the translations $D_1{}^\sigma$ of the subgroup D_1, and is therefore G-regular. In the second sequence (2) we have $H^r(G, K^*) = 0$ for all r. This fact depends only on the constant field $k_0{}^*$ being algebraicly closed. By Tsen's theorem we know that every element of k^* is a norm from K^*, which means that $H^0(G, K^*) = 0$. Coupling this with the fact that $H^1(G, K^*) = 0$, and noting that the same holds if G is replaced by any subgroup G_1 of G, we see that the G-module K^* has the property that its cohomology groups in two successive dimensions are trivial for all subgroups of G. From this it follows that they are trivial in all dimensions [7], [9]. For positive r the triviality of $H^r(G, K^*)$ has been noted by Nakayama and Hochschild [5]. In the third sequence (3) we have $H^r(G, k_0{}^*/U) = 0$ for all r, because $k_0{}^*/U$ has unique divisibility. Since k_0 is algebraicly closed, every element of $k_0{}^*$ has an n-th root, for every natural number n, and this n-th root is uniquely determined modulo U.

Using these cases of trivial cohomology groups in the cohomology sequences associated with our exact sequences of coefficient groups we obtain three isomorphisms:

$$\delta\colon H^r(G, E) \approx H^{r+1}(G, P), \tag{1'}$$

$$\delta\colon H^r(G, P) \approx H^{r+1}(G, k_0{}^*), \tag{2'}$$

$$i\colon H^r(G, U) \approx H^r(G, k_0{}^*). \tag{3'}$$

Putting these together we find finally:

$$i^{-1}\delta\delta\colon H^r(G, E) \approx H^{r+2}(G, U). \tag{4}$$

G operates simply on the group of roots of unity U, and U has a simple structure which depends only on the characteristic p of $k_0{}^*$. Thus in the situation we are considering—unramified extension and algebraicly closed constant field—the cohomology of G with coefficients in the divisor classes depends only on G and on the characteristic, not at all on the fields K and k.

The result we have obtained is remotely suggestive of class field theory. In class field theory, there is associated with each field K of a certain type a multiplicative group $A(K)$ such that, if G is a finite group of automorphisms of K, then G operates on $A(K)$ and we have $H^r(G, A(K)) \approx H^{r-2}(G, Z)$ for all r, where Z denotes the additive group of rational integers under trivial operation by G (cf. [6], [7], [9]). In local class field theory, $A(K)$ is the multiplicative group of the local field K. In global class field theory, $A(K)$ is the idèle class group of the number field K. In our present situation, where K is an unramified extension of the function field K, we can also find suitable coefficient groups $A(K)$ whose cohomology is like that of class field theory. To this effect we put $A(K) = E^*(K)$, where $E^*(K) = \mathrm{Hom}(E(K), U)$ is the group of all homomorphisms of the divisor class group of K into the roots of unity U. We call these homomorphisms characters of E and denote them by $\chi: E \to U$. If $\sigma: K \to K^\sigma$ is an isomorphism of K relative to k, we can define a corresponding isomorphism $\sigma: E^*(K) \to E^*(K^\sigma)$ by putting $\chi^\sigma(a^\sigma) = \chi(a)$ for $a \varepsilon E(K)$. The rule $(\chi^\sigma)^\tau = \chi^{(\tau\sigma)}$ is then obvious, and it follows that any group G of automorphisms of K/k operates in a natural way on $E^*(K)$.

To determine the cohomology groups $H^r(G, E^*(K))$ we need only use the duality theorem for the cohomology groups of finite groups which we now explain: Let W be a fixed abelian group on which G operates simply (W will be U in our applications). For any G-module X, let $X^* = \mathrm{Hom}(X, W)$. Then X^* and X are G-paired into W, and if $\phi \varepsilon H^r(G, X^*)$, the cup product multiplication by ϕ gives a map $\bar{\phi}$ of $H^{-1-r}(G, X)$ into $H^{-1}(G, W)$. Since $H^{-1}(G, W)$ can be identified with a subgroup of W, we may view this map $\bar{\phi}$ as an element of $(H^{-1-r}(G, X))^* = \mathrm{Hom}(H^{-1-r}(G, X), W)$. The duality theorem states that if W is divisible, then the map $\phi \to \bar{\phi}$ is an isomorphism: $H^r(G, X^*) \approx (H^{-1-r}(G, X))^*$; see, for example, [1] or [7].

In our field theoretic situation, U is divisible, because every root of unity has an n-th root for every natural number n, and this n-th root is again a root of unity. Therefore using (4) together with the duality theorem we find

$$H^r(G, E^*) \approx (H^{-1-r}(G, E))^* \approx (H^{1-r}(G, U))^*. \tag{5}$$

If Z is the additive group of integers with G operating simply, then the natural isomorphism $Z^* = \mathrm{Hom}(Z, U) \approx U$ is a G-isomorphism. Consequently, using the duality theorem again we find

$$H^{1-r}(G, U) \approx H^{1-r}(G, Z^*) \approx (H^{r-2}(G, Z))^*. \tag{6}$$

Combining (5) and (6) we obtain

$$H^r(G, E^*) \approx (H^{r-2}(G, Z))^{**}. \tag{7}$$

Now $H^{r-2}(G, Z)$ is a finite abelian group. If Y is any finite abelian group, the natural duality map $Y \to Y^{**}$ is an epimorphism (homomorphism onto). This follows in the usual way from the simple structure of U; namely, U contains at most one subgroup U_n of order n for each natural number n. In case the characteristic p is zero the map is an isomorphism, $Y \approx Y^{**}$, because then U_n exists for each n. In case $p > 0$, the map $Y \to Y^{**}$ has as kernel the p-primary part of Y (elements of p-power order), because then U_n exists only for n prime to p. Combining the natural map $H^{r-2}(G, Z) \to (H^{r-2}(G, Z))^{**}$ with the inverse of the isomorphism (7) we obtain therefore an epimorphism,

$$f: H^{r-2}(G, Z) \to H^r(G, E^*) \tag{8}$$

which is obviously canonical, and whose kernel is the p-primary part of $H^{r-2}(G, Z)$, p being the characteristic, and 0-primary part meaning the identity subgroup.

Having proved the existence of this epimorphism f, we can now describe it in more concrete terms. Reviewing the considerations of the preceding paragraphs, we see that f is simply the adjoint of the isomorphism (4)

$$i^{-1}\delta\delta: H^{-1-r}(G, E) \approx H^{1-r}(G, U),$$

with respect to the cup product duality. In other words, if $\zeta \in H^{r-2}(G, Z)$, then $f(\zeta) \in H^r(G, E^*)$ is uniquely determined by the property that

$$f(\zeta) \cdot \alpha = \zeta \cdot (i^{-1}\delta\delta\alpha), \tag{9}$$

for all $\alpha \in H^{-1-r}(G, E)$, where the dot denotes cup product multiplication. From this description of f, we can easily establish the rule $f(\zeta_1 \cdot \zeta) = f(\zeta_1) \cdot \zeta$ for any two integral cohomolgy classes ζ and ζ_1, if we use the fact that multiplication by an integral cohomology class ζ commutes with the natural cohomological operations i and δ. Namely, if α is any cohomology class over E of the appropriate dimension for determining $f(\zeta_1 \cdot \zeta)$ we have:

$$f(\zeta_1 \cdot \zeta) \cdot \alpha = \zeta_1 \cdot \zeta \cdot (i^{-1}\delta\delta\alpha) = \zeta_1 \cdot (i^{-1}\delta\delta(\zeta \cdot \alpha)) = f(\zeta_1) \cdot \zeta \cdot \alpha,$$

and since this is true for all α we can conclude $f(\zeta_1 \cdot \zeta) = f(\zeta_1) \cdot \zeta$, as stated.

Now putting $\zeta_1 = 1 \in H^0(G, Z)$, the unit element in the cohomology ring of G, and defining the fundamental two-dimensional class $\alpha^* \in H^2(G, E^*)$ by $\alpha^* = f(1)$, we see finally that $f(\zeta) = \alpha^* \cdot \zeta$. In other words, our canonical epimorphism (8), $H^{r-2}(G, Z) \to H^r(G, E^*)$, is obtained by multiplication with a fundamental two-dimensional class in E^*, just as in class field theory.

Our next task is to give an explicit characterization of the fundamental class $\alpha^* = f(1)$ by means of the standard 2-cocycles $\chi_{\sigma,\tau}$, which represent it.

Putting $\zeta = 1$ in (9) and applying the isomorphism i, we see that α^* is uniquely determined by the property

$$i(\alpha^* \cdot \alpha) = \delta\delta\alpha, \tag{10}$$

for all $\alpha \varepsilon H^{-3}(G, E)$. What we shall now do is simply write out formula (10) in terms of standard cochains. The fundamental class α^* will be represented by a standard 2-cocycle $\chi_{\sigma,\tau}$, which is a function $(\sigma, \tau) \rightarrow \chi_{\sigma,\tau}$ from $G \times G$ to E^* satisfying the identity $(\delta\chi)_{\rho,\sigma,\tau} = \chi^{\rho}{}_{\sigma,\tau}\chi^{-1}{}_{\rho\sigma,\tau}\chi_{\rho\sigma,\tau}\chi^{-1}{}_{\rho,\sigma} = 1$. Any (—3)-dimensional class α of G in E will be represented by a standard (—3)-cocycle $a_{\sigma,\tau}$, which is a function $(\sigma, \tau) \rightarrow a_{\sigma,\tau}$ from $G \times G$ to E satisfying the identity $(\delta a)_\tau = \prod_\sigma a^{\sigma^{-1}}{}_{\sigma,\tau} a^{-1}{}_{\sigma,\sigma^{-1}\tau} a_{\tau,\sigma} = 1$. The (—1)-dimensional class $\alpha^* \cdot \alpha$ is then described by the single element $\epsilon = \prod_{\sigma,\tau} \chi_{\sigma,\tau}(a_{\sigma,\tau})$, which is an n-th root of unity ($n = [K : k] =$ order of G). The left hand side of (10) is therefore the (—1)-dimensional class in $k_0{}^*$ determined by ϵ. To compute the right hand side of (10) we must choose divisors $\mathfrak{A}_{\sigma,\tau}$ in K representing the divisor classes $a_{\sigma,\tau}$. The coboundary of the (—3)-cochain $\mathfrak{A}_{\sigma,\tau}$ consists of principal divisors (A_τ), $A_\tau \varepsilon K^*$ such that

$$(A_\tau) = (\delta\mathfrak{A})_\tau = \prod_\sigma \mathfrak{A}^{\sigma^{-1}}{}_{\sigma,\tau}\mathfrak{A}^{-1}{}_{\sigma^{-1},\sigma\tau}\mathfrak{A}_{\tau,\sigma} \qquad (\text{in } D(K)). \tag{11}$$

The standard (—2)-cocycle (A_τ) then represents the (—2)-dimensional class $\delta\alpha \varepsilon H^{-2}(G, P)$, and the coboundary of the K^*-valued (—2)-cochain A_τ will be equal to the unique (—1)-cocycle ϵ which represents the class $\delta\delta\alpha \varepsilon H^{-1}(G, k_0{}^*)$:

$$\epsilon = \delta A = \prod_\tau A_\tau{}^{\tau^{-1}-1} \qquad (\text{in } K^*). \tag{12}$$

(ϵ is unique because G operates simply on $k_0{}^*$, so there are no (—1)-dimensional coboundaries.)

The standard two cocycles $\chi_{\sigma,\tau}$ of G in E^* which represent the fundamental class α^* are therefore those with the following property: Given any set of functions $A_\tau \varepsilon K^*$, and divisors $\mathfrak{A}_{\sigma,\tau} \varepsilon D(K)$ such that (11) holds, then the following equation holds:

$$\prod_{\sigma,\tau} \chi_{\sigma,\tau}(\mathfrak{A}_{\sigma,\tau}) = \prod_\tau A_\tau{}^{\tau^{-1}-1}. \tag{13}$$

Here the divisor class characters $\chi_{\sigma,\tau}$ are understood to be applied to divisors $\mathfrak{A}_{\sigma,\tau}$ in the obvious way.

To show that the parallel between our present situation and class field theory is complete, we must consider the interrelations among different fields. In class field theory, whenever a field k is a subfield of a field K, the associated module $A(k)$ is a submodule of $A(K)$; more precisely, there is a canonical monomorphism $i_{k/K}: A(k) \to A(K)$, which is usually viewed as an inclusion. In our case the role of this monomorphism is played by the conorm, $N^*_{k/K}: E^*(k) \to E^*(K)$, which is the dual of the norm, $N_{K/k}: E(K) \to E(k)$, and is defined by $(N^*_{k/K}\chi)(a) = \chi(N_{K/k}a)$ for $\chi \varepsilon E^*(k)$ and $a \varepsilon E(K)$. This conorm N^* is indeed a monomorphism, because the norm N is an epimorphism on divisor classes. In fact, the norm is an epimorphism even on the divisors themselves; a prime in the ground field is the norm of any one of the primes in the extension field dividing it, because the constant field is algebraicly closed. Of course the conorm monomorphism is transitive in the sense that $N^*_{k/K} = N^*_{k_1/K}N^*_{k/k_1}$, whenever $k \subset k_1 \subset K$.

There is on more property of the inclusion map $i_{k/K}: A(k) \to (K)$ which is crucial in class field theory. This can be expressed by saying that the fundamental theory of Galois Theory holds for the modules A. Precisely what we mean is this: If K/k is a normal extension with group $G = G(K/k)$, then the image of $A(k)$ in $A(K)$ consists exactly of the elements of $A(K)$ which are left fixed by the operation of G, or in a formula $i_{k/K}(A(k)) = (A(K))^G$. To show that the conorm has this property, i. e.,

$$N^*_{k/K}(E^*(k)) = (E^*(K))^G, \tag{14}$$

we note first that the left hand side is obviously included in the right. To show the converse, let $\chi \varepsilon (E^*(K))^G$. We wish to find a $\psi \varepsilon E^*(k)$ such that $\chi = N^*\psi$, that is, such that

$$\chi(a) = \psi(Na), \tag{15}$$

for all $a \varepsilon E(K)$. To this effect, we observe that the ψ we are looking for will be completely defined by (15) if it exists at all, because Na runs over the whole of $E(k)$ as a runs over $E(K)$. Therefore, we try to use (15) as a definition of ψ. This will work, provided we can show that $\chi(a) = 1$ whenever $Na = 1$, and this of course is where the assumption that χ is invariant under G must come in. Suppose therefore that $Na = 1$. Let $\mathfrak{A}$ be a divisor in K representing the divisor class a. Since $Na = 1$, $N\mathfrak{A}$ is principal, and there exists a function $\alpha \varepsilon k$ such that $N\mathfrak{A} = (\alpha)$. Since $H^0(G, K^*) = 0$ there exists a function $A \varepsilon K$ such that $NA = \alpha$. It follows that $N(\mathfrak{A}A^{-1}) = (1)$, and consequently there exist divisors $\mathfrak{B}_\sigma$ in K, $\sigma \varepsilon G$, such that

$$\mathfrak{A}A^{-1} = \prod_\sigma \mathfrak{B}_\sigma^{\sigma-1}.$$

(This follows from $H^{-1}(G, D) = 0$, but can easily be seen directly if we recall that K/k is assumed unramified.) Finally, letting $b_\sigma \varepsilon E(K)$ be the divisor class of the divisor $\mathfrak{B}_\sigma$ and going over to divisor classes, we find

$$a = \prod_\sigma b_\sigma{}^{\sigma-1}.$$

Since χ is invariant under G we have $\chi(b_\sigma{}^\sigma) = \chi(b_\sigma)$ for all σ and consequently $\chi(a) = 1$, which is what we were trying to prove.

There is of course a map of divisor class characters which goes in the opposite direction to the conorm. We call it the norm and denote it by $N_{K/k}: E^*(K) \to E^*(k)$. We define this norm as the dual of the natural injection $i_{k/K}: E(k) \to E(K)$ of divisor classes of the subfield k into those of the extension K, that is,

$$(N\chi)(a) = \chi(ia), \tag{16}$$

for $\chi \varepsilon E^*(K)$ and $a \varepsilon E(k)$. The justification for the term norm for this map rests in the formula

$$N^*(N\chi) = \prod_\sigma \chi^\sigma, \tag{17}$$

which states that $N\chi$ is essentially—i. e., when imbedded in $E^*(K)$ by the conorm becomes—the product of the conjugates of χ. To prove (17) we simply write, for $a \varepsilon E(K)$,

$$\begin{aligned}(N^*(N\chi))(a) &= (N\chi)(Na) = \chi(iNa) = \chi(\prod_\sigma a^{\sigma^{-1}}) \\ &= \prod_\sigma \chi(a^{\sigma^{-1}}) = \prod_\sigma \chi^\sigma(a) = (\prod_\sigma \chi^\sigma)(a).\end{aligned}$$

From (14) and (17) we see that the conorm N^* induces a natural isomorphism of the group of divisor class character norm residues of the ground field k, $E^*(k)/N(E^*(K))$, onto the factor group of divisor class characters of K which are invariant under G, modulo those which are products of conjugates. This latter factor group may of course be identified with the 0-dimensional cohomology group $H^0(G, E^*(K))$. Therefore we have

$$H^0(G, E^*(K)) \approx E^*(k)/N(E^*(K)). \tag{18}$$

If K/k is abelian, of degree $n = [K:k]$ prime to p, and if we use the natural isomorphism

$$H^{-2}(G, Z) \approx G, \tag{19}$$

then we can interpret the case $r = 0$ of our natural epimorphism (8) as an isomorphism

$$f: G \approx E^*(k)/N(E^*(K)) \tag{20}$$

(isomorphism because of the assumption that G has no elements of order p). This isomorphism f is the analog of the reciprocity law isomorphism of class field theory. In class field theory, an explicit formula for this isomorphism is given by the Nakayama map associated with a fundamental 2-cocycle. The same holds in our present case, because the Nakayama map is simply an explicit formula for the cup product of an element of $H^{-2}(G,Z)$ which corresponds to $\rho \varepsilon G$ under the isomorphism (19), with the fundamental class, and our map f is defined by such a product. Therefore, as an explicit formula for (20) we have, for $\rho \varepsilon G$,

$$f(\rho) = \psi_\rho \quad (\operatorname{mod} N(E^*(K))), \tag{21}$$

if $\psi_\rho \varepsilon E^*(k)$ is defined by

$$N^*\psi_\rho = \prod_\sigma \chi_{\sigma,\rho}, \tag{22}$$

$\chi_{\sigma,\rho}$ being a fundamental 2-cocycle.

The norm group $N(E^*(K))$ can easily be described directly; it consists of those characters ψ of $E^*(k)$ which are trivial on the kernel of the injection map $i: E(k) \to E(K)$, i. e., which satisfy the condition $\psi(i^{-1}(1)) = 1$. Indeed, this condition is obviously necessary and sufficient for the existence of a character χ_0 on the subgroup $i(E(k))$ of $E(K)$ such that $\psi = \chi_0 i$; and since any such χ_0 can be extended to a character of $E(K)$ with the aid of Zorn's lemma and the divisibility of U, we see that our condition is necessary and sufficient for the existence of $\chi \varepsilon E^*(K)$, such that $\psi = \chi i$, i. e., $\psi = N\chi$.

From what we have said it follows that to determine the coset of a character $\psi \varepsilon E^*(k)$ modulo the norm group $N(E^*(K))$ is simply to give the values of ψ on the subgroup $i^{-1}(1)$ of $E(k)$. This shows what we must do in order to give an explicit formula for the reciprocity law isomorphism (20): for each $\rho \varepsilon G$ and each $c \varepsilon i^{-1}(1)$ we must calculate the root of unity $\psi_\rho(c)$, where ψ_ρ is defined by equation (22), $\chi_{\sigma,\rho}$ being a fundamental 2-cocycle. The result is the following: Let $\mathfrak{c}$ be a divisor in k representing the class of c. Since $ic = 1$, $\mathfrak{c}$ becomes principal in K, i. e., $\mathfrak{c} = (B)$ with $B \varepsilon K^*$. Then

$$\psi_\rho(c) = B^{\rho^{-1}-1}. \tag{23}$$

To prove this, let $\mathfrak{B}$ be a divisor in K such that $\mathfrak{c} = N\mathfrak{B}$. In K this equation reads

$$(B) = \prod_\sigma \mathfrak{B}^{\sigma^{-1}}.$$

This last equation can be put in the form of equation (11) if we define the (-1)-cochain $A_\tau \varepsilon K^*$ and the (-2)-cochain $\mathfrak{A}_{\sigma,\tau} \varepsilon D(K)$ by $A_\tau = 1$ or

$A_\tau = B$ according as $\tau \neq \rho$ or $\tau = \rho$ and $\mathfrak{A}_{\sigma,\tau} = 1$ or $\mathfrak{A}_{\sigma,\tau} = \mathfrak{B}$ according as $\tau \neq \rho$ or $\tau = \rho$. The resulting equation (13) then reads

$$\prod_{\sigma} \chi_{\sigma,\rho}(\mathfrak{B}) = B^{\rho^{-1}-1}.$$

Comparing this with (22) and noting $N\mathfrak{B} = \mathfrak{c} \varepsilon c$, we see that (23) holds as contended.

Of course the proper interpretation of the result which we have obtained is simply to view it as giving the relationship between our pseudo class field theory of unramified abelian extensions of degree prime to p, and the Kummer theory of such extensions. Using the Kummer theory, one easily proves the appropriate existence theorem for our pseudo class field theory, namely, a subgroup X of $E^*(k)$ is the norm subgroup of some class field K/k if and only if there exists a subgroup Y of $E(k)$ of finite order n prime to p, such that X consists of all characters trivial on Y. The class field is then obtained by taking for each element of Y a representative divisor $\mathfrak{c}$, and then extracting the n-th roots of elements $\gamma \varepsilon k^*$ for which $(\gamma) = \mathfrak{c}^n$.

In order to complete the discussion of the parallel between our situation and class field theory, we must consider the relations between the fundamental 2-dimensional classes in different extensions. For these considerations we drop the assumption that the extensions are abelian.

Let $\bar{k} \subset k \subset K$ with $K/\bar{k}$ normal unramified, but $k/\bar{k}$ not necessarily normal. We wish to prove that the restriction to K/k of the fundamental class $\tilde{\alpha}^*$ of $K/\bar{k}$ is equal to the fundamental class α^* of K/k. For the proof we let $\tilde{G} = G(K/\bar{k})$ and $G = G(K/k)$, so that G is a subgroup of $\tilde{G}$. Let elements $A_\tau \varepsilon K^*$ and divisors $\mathfrak{A}_{\sigma,\tau} \varepsilon D(K)$ be given satisfying equation (11) for $\sigma, \tau \varepsilon G$. Extend the definition of these cochains A and $\mathfrak{A}$ to the big group $\tilde{G}$ by defining $A_\tau = 1$ for $\tau \not\varepsilon G$ and $\mathfrak{A}_{\sigma,\tau} = (1)$ for σ or $\tau \not\varepsilon G$. Then equation (12) still holds if G is replaced by $\tilde{G}$, i. e., if the product is taken over all of $\tilde{G}$ instead of over the subgroup G. If $\chi = \chi_{\sigma,\tau}$ is a fundamental 2-cocycle for the extension $K/\bar{k}$, then we obtain from (13)

$$\prod_{\sigma,\tau \varepsilon G} \chi_{\sigma,\tau}(\mathfrak{A}_{\sigma,\tau}) = \prod_{\sigma,\tau \varepsilon \tilde{G}} \chi_{\sigma,\tau}(\mathfrak{A}_{\sigma,\tau}) = \prod_{\tau \varepsilon \tilde{G}} {A_\tau}^{\tau^{-1}-1} = \prod_{\tau \varepsilon G} {A_\tau}^{\tau^{-1}-1}.$$

This shows that the restriction of χ to the subgroup G is a fundamental 2-cocycle for the extension K/k as contended.

Next, suppose $k \subset \bar{K} \subset K$ with both extensions normal and unramified. Let $m = [K:\bar{K}]$. We wish to prove that the inflation of the fundamental class of $\bar{K}/k$ is equal to the m-th power of the fundamental class of K/k. Let $\tilde{G} = G(\bar{K}/k)$ and $G = G(K/k)$ so that $\tilde{G}$ is a factor group of G. Let functions A_τ and divisors $\mathfrak{A}_{\sigma,\tau}$ be given satisfying equation (11) for $\sigma, \tau \varepsilon G$. Viewing the elements of $\bar{G}$ as cosets in G, and denoting them by s, t, etc., define

$$\bar{A}_t = \prod_{\tau \varepsilon t} N_{K/\bar{K}} A_\tau ; \qquad \bar{\mathfrak{A}}_{s,t} = \prod_{\sigma \varepsilon s,\, \tau \varepsilon t} N_{K/\bar{K}} \mathfrak{A}_{\sigma,\tau}.$$

Then from the fact that $\bar{K}/k$ is normal, one sees easily that the quantities $\bar{A}_t$ and $\mathfrak{A}_{s,t}$ satisfy the condition (11) for the extension $\bar{K}/k$, and furthermore,

$$(24) \qquad \prod_t \bar{A}_t{}^{t^{-1}-1} = N_{K/\bar{K}}(\prod_\tau A_\tau{}^{\tau^{-1}-1}) = (\prod_\tau A_\tau{}^{\tau^{-1}-1})^m,$$

because $\prod_\tau A_\tau{}^{\tau^{-1}-1} \varepsilon k_0{}^*$. Consequently, if $\bar{\chi} = \bar{\chi}_{s,t}$ is a fundamental 2-cocycle for $\bar{K}/k$ we obtain from (13) and (24)

$$\begin{aligned}\prod_\tau (A_\tau{}^{\tau^{-1}-1})^m &= \prod_{s,t} \bar{\chi}_{s,t}(\bar{\mathfrak{A}}_{s,t}) = \prod_{s,t} \prod_{\sigma \varepsilon s,\, \tau \varepsilon t} \bar{\chi}_{s,t}(N_{K/\bar{K}} \mathfrak{A}_{\sigma,\tau}) \\ &= \prod_{s.t} \prod_{\sigma \varepsilon s,\, \tau \varepsilon t} (N^*{}_{\bar{K}/K} \bar{\chi}_{s,t})(\mathfrak{A}_{\sigma,\tau}) = \prod_{\sigma,\tau \varepsilon G} (\operatorname{infl} \bar{\chi})_{\sigma,\tau}(\mathfrak{A}_{\sigma,\tau}),\end{aligned}$$

where $\operatorname{infl} \bar{\chi}$ is the inflation to G of $\bar{\chi}$. Since the m-th power of the usual left hand side occurs in this equation, we conclude that $\operatorname{infl} \bar{\chi}$ represents the m-th power of the fundamental class of K/k as contended.

II. The Classical Case.

We turn now to the classical case in which the constant field k_0 is the field of complex numbers. In addition to our old notations which will still be used in the present section we introduce the following new ones:

$D_0(K)$ = the group of divisors of degree 0 of K;

$E_0(K)$ = the group of divisor classes of degree 0 of K;

T = the group of unimodular complex numbers (one dimensional torus group);

$g(K)$ = genus of K;

$R(K)$ = the Riemann surface of K;

$H_1(K) = H_1(R(K), Z)$ = the 1-dimensional homology group of $R(K)$ with integer coefficients.

As is well known, $H_1 \approx Z^{2g}$ is a free abelian group on $2g$ generators, $E_0 \approx T^{2g}$ is a $2g$-dimensional torus group, and there is a natural duality between H_1 and E_0 in the sense of Pontrjagin. Giving E_0 the compact topology of T^{2g}, we can extend this topology to E in a unique way such that the factor group $E/E_0 \approx Z$ is discrete, i. e., such that E_0 is an open subgroup of E. Then E is a locally compact group and from the exact sequence

$$(1) \qquad \{1\} \to E_0 \to E \to Z \to \{1\},$$

we obtain the dual exact sequence

$$(2) \qquad \{1\} \leftarrow H_1 \leftarrow \hat{E} \leftarrow T \leftarrow \{1\},$$

which shows that the Pontrjagin character group $\hat{E}$ of E has an open compact subgroup T, with the factor group isomorphic to the discrete group $H_1 \approx Z^{2g}$. Of course the exact sequences (1) and (2) split; if a_1 is any fixed divisor class of degree 1, then any divisor class can be written uniquely in the form aa_1^{ν} with $a \in E_0$, $\nu \in Z$, so we have $E \approx E_0 \times Z \approx T^{2g} \times Z$ and accordingly $\hat{E} \approx H_1 \times T \approx Z^{2g} \times T$. These direct product decompositions are topological as well as algebraic, but they are not canonical since they depend on the choice of a_1.

The advantage of the classical case, which we are now discussing over the algebraic case discussed in the preceding section, is that we can use the topological character group $\hat{E}$, whose simple structure we have just described, in place of the algebraic character group E^*. E^* was defined to be the group of all homomorphisms of E into the group of roots of unity U; $E^* = \mathrm{Hom}(E, U)$. $\hat{E}$ is the group of continuous homomorphisms of E into T; $\hat{E} = \mathrm{Cont\,Hom}(E, T)$. Both these groups are subgroups of the group $\mathrm{Hom}(E, T)$ of all homomorphisms of E into T. The first thing to remark is that all the facts we have proven in **I** about $E^* = \mathrm{Hom}(E, U)$ are true for $\mathrm{Hom}(E, T)$ as well, because the only properties of U which we used in **I** are possessed also by T, namely: (i) k_0^*/T has unique divisibility; (ii) T is divisible; and (iii) T has exactly one subgroup of order n for each natural number n. Next, in order to show that the cohomology groups with coefficients in $\mathrm{Hom}(E, T)$ are the same as those with coefficients in $\hat{E} = \mathrm{Cont\,Hom}(E, T)$, we will prove that the factor group $\mathrm{Hom}(E, T)/\mathrm{Cont\,Hom}(E, T)$ has unique divisibility. Since $E \approx T^{2g} \times Z$ is a direct product of T's and a Z, it suffices to prove the statement for the factors T and Z separately instead of for E itself. Z is no problem; being discrete, its homomorphisms are the same as its continuous homomorphisms. Concerning T we must prove: (a) For each natural number n, the homomorphisms $f: T \to T$ can be written in the form $f = g^n\phi$, where ϕ is continuous (divisibility of the factor group $\mathrm{Hom}(T, T)/\mathrm{Cont\,Hom}(T, T)$); and (b) If the n-th power of a homomorphism $g: T \to T$ is continuous, then g itself is continuous (uniqueness of the divisibility).

To prove (a), consider the group U_n of n-th roots of unity. A homomorphism $f: T \to T$ carries U_n into itself, and since U_n is cyclic, there exists an integer m such that $f(\epsilon) = \epsilon^m$ for $\epsilon \in U_n$. Define $\phi(\lambda) = \lambda^m$ for all $\lambda \in T$. Then $f(\epsilon) = \phi(\epsilon)$ for $\epsilon \in U_n$, so we can define a homomorphism $g: T \to T$ by $g(\lambda) = f(\sqrt[n]{\lambda})/\phi(\sqrt[n]{\lambda})$, this last quantity being independent of the choice of the n-th root. Hence $g^n(\lambda) = g(\lambda^n) = f(\lambda)/\phi(\lambda)$, or $f = g^n\phi$ as contended. Now consider statement (b). If g^n is continuous, then $g^n(\lambda) = \lambda^m$ for some integer m, because the only continuous homomorphisms of T into T are of this form. For $\epsilon \in U_n$ we find $g^n(\epsilon) = 1 = \epsilon^m$. Hence n divides m, say $m = nl$. It follows for any $\lambda \in T$ that $g(\lambda^n) = g^n(\lambda) = (\lambda^n)^l$, and since $T^n = T$, this implies $g(\lambda) = \lambda^l$ for all λ. Therefore g is continuous as contended.

Suppose now that our field K is a finite normal unramified extension of k with Galois group $G = G(K/k)$. Then, according to the remarks we have just made, we obtain an isomorphism

$$f: H^{r-2}(G, Z) \approx H^r(G, \hat{E}(K)) \tag{3}$$

of the same type as the isomorphism (8) of part I. Just as in the algebraic case, this isomorphism is given by the cup product with the canonical generator of $H^2(G, \hat{E})$. Soon we shall give a topologico-analytic description of this canonical 2-dimensional class. But first let us make a few remarks about the cohomology of G with coefficients in the homology group $H_1(K)$. The exact sequence (2) gives rise to the exact cohomology sequence

$$\cdots \to H^r(G, \hat{E}) \to H^r(G, H_1) \to H^{r+1}(G, T) \to H^{r+1}(G, \hat{E}) \to \cdots. \tag{4}$$

Using (3) together with $\delta: H^r(G, T) \approx H^{r+1}(G, Z)$, we can reinterpret (4) as

$$\cdots \to H^{r-2}(G, Z) \to H^r(G, H_1) \to H^{r+2}(G, Z) \to H^{r-1}(G, Z) \to \cdots. \tag{5}$$

This exact sequence (5) gives us some information about the groups $H^r(G, H_1)$. However, the problem of $H^r(G, H_1)$ can be viewed as a purely topological problem, and we should point out that an exact sequence of the same type as (5) can be obtained by standard topological methods. To this effect we forget about the analytic structure of the Riemann surfaces and simply view $R(K)$ as a finite regular unramified covering of $R(k)$ with covering transformation group G. By taking a cellular decomposition of $R(k)$ and lifting it up to the covering surface $R(K)$, we obtain a decomposition of $R(K)$ in which the cells are permuted by G, no cell being left fixed by any transformation other than the identity. Therefore in the associated chain complex

$$\cdots \to 0 \to C_2 \xrightarrow{\partial} C_1 \xrightarrow{\partial} C_0 \to 0 \to \cdots, \tag{6}$$

the chain groups C_r are free G-modules; and of course the boundary homomorphism δ is a G-homomorphism. Letting Z_j, B_j, and H_j denote the groups of j-cycles, j-boundaries, and j-homology classes in this complex, we can break the complex up into four exact sequences:

$$0 \to H_2 \xrightarrow{i} C_2 \xrightarrow{\partial} B_1 \to 0, \tag{7}$$

$$0 \to B_1 \xrightarrow{i} Z_1 \to H_1 \to 0, \tag{8}$$

$$0 \to Z_1 \xrightarrow{i} C_1 \xrightarrow{\partial} B_0 \to 0, \tag{9}$$

$$0 \to B_0 \xrightarrow{i} C_0 \to H_0 \to 0, \tag{10}$$

where i is the inclusion map. Now consider the cohomology groups associated with the operation of G on these modules. Since the C_j are G-free, and G is finite, we have $H^r(G, C_j) = 0$ for $j = 0, 1, 2$. Since $H_2 \approx Z \approx H_0$ we obtain from (7), and from (9) and (10), the isomorphisms

$$\delta_{(7)} : H^{r-1}(G, B_1) \approx H^r(G, H_2) \approx H^r(G, Z),$$

$$\delta_{(9)}\delta_{(10)} : H^{r-2}(G, Z) \approx H^{r-2}(G, H_0) \approx H^r(G, Z_1).$$

If we insert these results about B_1 and Z_1 in the following exact cohomology sequence associated with (8),

$$\cdots \to H^r(G, Z_1) \to H^r(G, H_1) \to H^{r+1}(G, B_1) \to H^{r+1}(G, Z_1) \to \cdots, \tag{11}$$

then we obtain an exact sequence which is of exactly the same form as (5), and which is presumably identical with (5), save possibly for the signs of the mappings. However, since this topological approach is a digression from our main line of thought, we will not enter further into the matter here.

In order to describe explicitly the canonical duality between H_1 and E_0 which we have used in the discussion above, we will make use of certain abelian differentials of the third kind which are attached to divisors of degree 0. If $\mathfrak{A} = \prod \mathfrak{P}^{\nu_{\mathfrak{P}}}$, $\sum \nu_{\mathfrak{P}} = 0$, is a divisor of degree 0, then the corresponding differential is denoted symbolically by $d \log \mathfrak{A}$ and is uniquely characterized by the following two properties: (i) At each point $\mathfrak{P} \varepsilon R$, $d \log \mathfrak{A}$ has a simple pole with residue $\nu_{\mathfrak{P}} =$ order of $\mathfrak{A}$ at $\mathfrak{P}$ (a simple pole with residue 0 is no pole at all!) (ii) All periods of $d \log \mathfrak{A}$ are pure imaginary. One method of showing the existence of these differentials is the following. The Riemann-Roch theorem guarantees the existence of a differential satisfying condition (i), because $\sum \nu_{\mathfrak{P}} = 0$. Its local periods are then integral multiples of $2\pi i$, so condition (ii) is now merely a condition on the global periods. Since there exists a unique differential of first kind whose $2g$ basic global periods have arbitrarily prescribed real parts, we can correct our original differential of the third kind by one of the first so that condition (ii) is satisfied. From the characterization of these differentials by properties, it is clear that the rule

$$d \log (\mathfrak{A}\mathfrak{B}) = d \log \mathfrak{A} + d \log \mathfrak{B} \qquad (\mathfrak{A}, \mathfrak{B} \,\varepsilon\, D_0) \tag{12}$$

holds, and furthermore, if $\mathfrak{A} = (A)$ is the divisor of a function A, then

$$d \log \mathfrak{A} = d \log A = dA/A. \tag{13}$$

We appeal to (12) and (13) for justification of our somewhat unconventional notation $d \log \mathfrak{A}$.

If B is a 1-chain on the surface R, and $\mathfrak{A}$ a divisor of degree 0, none of whose poles or zeros lie on the boundary of B, we define the symbol

$$(\mathrm{B}, \mathfrak{A}) = \exp \int_{\mathrm{B}'} d \log \mathfrak{A}, \tag{14}$$

where B' is any 1-chain homologous to B which avoids the singularities of $d \log \mathfrak{A}$. B' can be obtained by deforming the cells of B slightly, keeping their end points fixed. Since the local periods of $d \log \mathfrak{A}$ are integral multiples of $2\pi i$, the right side of (14) is independent of the choice of B'. The following rules for our symbol are evident from the definition

$$(\mathrm{B}_1 + \mathrm{B}_2, \mathfrak{A}) = (\mathrm{B}_1, \mathfrak{A}) (\mathrm{B}_2, \mathfrak{A}), \tag{15}$$

$$(\mathrm{B}, \mathfrak{A}_1\mathfrak{A}_2) = (\mathrm{B}, \mathfrak{A}_1) (\mathrm{B}, \mathfrak{A}_2). \tag{16}$$

Furthermore, if σ is a one-to-one conformal map of R on R, then we have

$$(\sigma\mathrm{B}, \mathfrak{A}^{\sigma}) = (\mathrm{B}, \mathfrak{A}), \tag{17}$$

because σ viewed as an integral substitution changes B into $\sigma\mathrm{B}$ and $d \log A$ into a differential with the characterizing properties of $d \log (\mathfrak{A}^{\sigma})$, hence into $d \log (\mathfrak{A}^{\sigma})$. Finally, if $A \,\varepsilon\, K^*$ is a function and the boundary of B is $\partial\mathrm{B} = \sum \nu_{\mathfrak{Q}} \mathfrak{Q}$, then we see from (13) that

$$(\mathrm{B}, (A)) = A \mid_{\partial\mathrm{B}} = \prod_{\mathfrak{Q}} (A(\mathfrak{Q}))^{\nu_{\mathfrak{Q}}}. \tag{18}$$

Let us now consider the symbol $(\Gamma, \mathfrak{A})$ for a 1-cycle, Γ. It is defined for all $\mathfrak{A} \,\varepsilon\, D_0$ because $\partial\Gamma$ is empty. Of course it depends only on the homology class $\bar{\Gamma}$ of Γ, and from (18) we see that it depends only on the divisor class $\mathfrak{A}$ of $\mathfrak{A}$. Since the periods of $d \log \mathfrak{A}$ are pure imaginary $(\Gamma, \mathfrak{A})$ has absolute value 1, i. e., $(\Gamma, \mathfrak{A}) \,\varepsilon\, T$. Therefore the symbol $(\bar{\Gamma}, \bar{\mathfrak{A}}) = (\Gamma, \mathfrak{A})$ is a pairing of H_1 and E_0 into T. In fact,, the discrete group H_1 and the compact group E_0 are dual under this pairing (see Weyl [2] or Igusa [3]). In the proof of the duality it is necessary to use the so-called theorem of interchange of argument and parameter to obtain a formula for $(\Gamma, \mathfrak{A})$ in terms of differentials of the first kind. Namely, if $\mathfrak{A} = \mathfrak{P}_1\mathfrak{P}_2 \cdots \mathfrak{P}_m\mathfrak{Q}_1^{-1}, \mathfrak{Q}_2^{-1} \cdots \mathfrak{Q}_m^{-1}$ is any divisor of degree 0, the interchange shows that

$$(\Gamma, \mathfrak{A}) = \exp 2\pi i \mathcal{R} \{ \sum_{j=1}^{m} \int_{\mathfrak{Q}_j}^{\mathfrak{P}_j} dw_\Gamma \},$$

where dw_Γ is the uniquely determined differential of the first kind such that

$$\mathcal{R} \int_{\Gamma'} dw_\Gamma = \text{intersection number of } \Gamma' \text{ and } \Gamma \text{ for all cycles } \Gamma'.$$

We can now describe the divisor class characters $\chi \varepsilon \hat{E}$ explicitly. As usual, we will sometimes view χ as a divisor character which is trivial on principal divisors, writing $\chi(\mathfrak{A}) = \chi(\bar{\mathfrak{A}})$ if $\bar{\mathfrak{A}}$ is the class of the divisor $\mathfrak{A}$. Let us select a prime divisor $\mathfrak{P}$, to be kept fixed during the rest of the discussion, so that an arbitrary divisor can be written uniquely in the form $\mathfrak{A}\mathfrak{P}^s$ with $\mathfrak{A} \varepsilon D_0$, $s \varepsilon Z$. Then if Γ is any 1-cycle, and λ any complex number of absolute value 1, we can define a character χ by the formula

$$\chi(\mathfrak{A}\mathfrak{P}^s) = (\Gamma, \mathfrak{A})\lambda^s. \tag{19}$$

We shall use the notation $\chi = \chi(\Gamma, \lambda)$ to refer to this character. Every character is of this form. The number $\lambda = \chi(\mathfrak{P})$ is uniquely determined by χ, as is the homology class $\bar{\Gamma}$ of Γ. The maps in the exact sequence (2), $T \rightarrow \hat{E} \rightarrow H_1$, are given by $\lambda \rightarrow \chi(0, \lambda)$ and $\chi(\Gamma, \lambda) \rightarrow \bar{\Gamma}$.

Suppose now that our field K is a normal unramified extension of k with Galois group $G = G(K/k)$. We contend that the automorphisms $\rho \varepsilon G$ operate on characters according to the rule

$$(\chi(\Gamma, \lambda))^\rho = \chi(\rho\Gamma, \lambda(\Gamma, \mathfrak{P}^{\rho^{-1}-1})). \tag{20}$$

Indeed from (17) we see that both sides of (20) yield the same value when applied to $\mathfrak{A} \varepsilon D_0$. The remaining computation

$$\chi^\rho(\mathfrak{P}) = \chi(\mathfrak{P}^{\rho^{-1}}) = \chi(\mathfrak{P})\chi(\mathfrak{P}^{\rho^{-1}-1}) = \lambda(\Gamma, \mathfrak{P}^{\rho^{-1}-1}),$$

shows that (20) is correct.

We can now give an explicit description of a fundamental 2-cocycle $\chi_{\sigma,\tau}$. Let $\mathfrak{Q}$ be a fixed base point on the surface $R(K)$, none of whose conjugates $\mathfrak{Q}^\sigma$, $\sigma \varepsilon G$, coincides with the point $\mathfrak{P}$. For each $\sigma \varepsilon G$, select a fixed path B_σ running from $\mathfrak{Q}$ to $\mathfrak{Q}^\sigma$. Then for each pair of elements $\sigma, \tau \varepsilon G$ the chain

$$\Gamma_{\sigma,\tau} = \mathrm{B}_\sigma + \sigma\mathrm{B}_\tau - \mathrm{B}_{\sigma\tau} \tag{21}$$

is a cycle, and as function of σ, τ, $\Gamma_{\sigma,\tau}$ is a standard 2-cocycle of G in the group of cycles. We contend that by suitable choice of numbers $\lambda_{\sigma,\tau} \varepsilon T$ we will obtain a fundamental 2-cocycle of the form

$$(22) \qquad \chi_{\sigma,\tau} = \chi(\Gamma_{\sigma,\tau}, \lambda_{\sigma,\tau}).$$

Leaving the choice of the numbers $\lambda_{\sigma,\tau}$ open for the moment, we compute the coboundary of $\chi_{\sigma,\tau}$ by means of (20), obtaining

$$(\delta\chi)_{\rho,\sigma,\tau} = ((\delta\Gamma)_{\rho,\sigma,\tau}, (\delta\lambda)_{\rho,\sigma,\tau}(\Gamma_{\sigma,\tau}, \mathfrak{P}^{\rho^{-1}-1})).$$

Since $\delta\Gamma = 0$, the condition on the λ's for $\chi_{\sigma,\tau}$ to be a cocycle is simply

$$(23) \qquad (\delta\lambda)_{\rho,\sigma,\tau} = (\Gamma_{\sigma,\tau}, \mathfrak{P}^{\rho^{-1}-1})^{-1}.$$

To define the numbers $\lambda_{\sigma,\tau}$ we use the following notation: if z is any non-zero complex number, we put $\text{u.p.}z = z/|z|$ and call this the "unimodular part" of z. Putting

$$(24) \qquad \lambda_{\sigma,\tau} = \text{u.p.}(\mathrm{B}_\tau, \mathfrak{P}^{\rho^{-1}-1}) = \exp\left(i\vartheta \int_{\mathrm{B}_\tau} d\log(\mathfrak{P}^{\sigma^{-1}-1})\right),$$

we find

$$(\delta\lambda)_{\rho,\sigma,\tau} = \text{u.p.}(\mathrm{B}_\tau, \mathfrak{P}^{\sigma^{-1}-1})(\mathrm{B}_\tau, \mathfrak{P}^{(\rho\sigma)^{-1}-1})^{-1}(\mathrm{B}_{\sigma\tau}, \mathfrak{P}^{\rho^{-1}-1})(\mathrm{B}_\sigma, \mathfrak{P}^{\rho^{-1}-1})^{-1}.$$

Combining the first two terms on the right and using rule (17) gives

$$\begin{aligned}(\delta\lambda)_{\rho,\sigma,\tau} &= \text{u.p.}(\sigma\mathrm{B}_\tau, \mathfrak{P}^{\rho^{-1}-1})^{-1} \quad (\mathrm{B}_{\sigma,\tau}, \mathfrak{P}^{\rho^{-1}-1})(\mathrm{B}_\sigma, \mathfrak{P}^{\rho^{-1}-1})^{-1} \\ &= \text{u.p.}(\Gamma_{\sigma,\tau}, \mathfrak{P}^{\rho^{-1}-1})^{-1}.\end{aligned}$$

Since $\Gamma_{\sigma,\tau}$ is a cycle, a symbol of the form $(\Gamma_{\sigma,\tau}, \mathfrak{A})$ is unimodular, and we can drop the u.p. Therefore (23) is satisfied, and our 2-cochain $\chi_{\sigma,\tau}$ is a 2-cocycle.

It is a straightforward matter to check by means of the criterion contained in formulas (11) and (13) of the first half of this paper that $\chi_{\sigma,\tau}$ is a fundamental 2-cocycle. We write the divisor (-3)-cochain in the form $\mathfrak{A}_{\sigma,\tau}\mathfrak{P}^{s_{\sigma,\tau}}$ with $\mathfrak{A}_{\sigma,\tau} \varepsilon D_0$, $s_{\sigma,\tau} \varepsilon Z$. Denoting the function (-2)-cochain as usual by A_τ, equation (I, 11) takes the form

$$(25) \qquad (A_\tau) = \prod_{\sigma \varepsilon G} \mathfrak{A}^{\sigma^{-1}}{}_{\sigma,\tau}\mathfrak{A}^{-1}{}_{\sigma^{-1},\sigma\tau}\mathfrak{A}_{\tau,\sigma}\mathfrak{P}^{\kappa},$$

where κ stands for the (unprintable) exponent $\sigma^{-1}s_{\sigma,\tau} - s_{\sigma^{-1},\sigma\tau} + s_{\tau,\sigma}$. Assuming this, we must then show

$$(26) \qquad \prod_{\sigma,\tau} \chi_{\sigma,\tau}(\mathfrak{A}_{\sigma,\tau}\mathfrak{P}^{s_{\sigma,\tau}}) = \prod_{\tau} A_\tau{}^{\tau^{-1}-1}.$$

After two preliminary remarks, it will turn out that (26) follows from a straightforward computation. First, taking the degrees of the divisors equated in (25) we find $\sum_{\sigma}(s_{\sigma,\tau}-s_{\sigma^{-1},\sigma\tau}+s_{\tau,\sigma})=0$. Subtracting this quantity from the exponent κ of $\mathfrak{P}$ in (25) we obtain (25) in the simpler form:

$$(A_\tau)=\prod_{\sigma\varepsilon G}\mathfrak{A}^{\sigma^{-1}}{}_{\sigma,\tau}\mathfrak{A}^{-1}{}_{\sigma^{-1},\sigma\tau}\mathfrak{A}_{\tau,\sigma}\mathfrak{P}^{(\sigma^{-1}-1)s_{\sigma,\tau}}. \tag{25'}$$

Secondly, we remark that it is no loss of generality to assume that the divisors $\mathfrak{A}_{\sigma,\tau}$ have no zeros or poles at the places $\mathfrak{O}^\sigma$, $\sigma\,\varepsilon\,G$, so that the symbols of type $(\mathrm{B}_\rho,\mathfrak{A}_{\sigma,\tau})$ are defined. This follows from the fact that the (-3)-cochain $\mathfrak{A}_{\sigma,\tau}\mathfrak{P}^{s_{\sigma,\tau}}$ serves to represent a divisor class (-3)-cocycle, and consequently each $\mathfrak{A}_{\sigma,\tau}$ may be changed by an arbitrary principal divisor.

Now to prove (26) we write

$$\prod_{\sigma,\tau}\chi_{\sigma,\tau}(\mathfrak{A}_{\sigma,\tau}\mathfrak{P}^{s_{\sigma,\tau}})=\prod_{\sigma,\tau}(\Gamma_{\sigma,\tau},\mathfrak{A}_{\sigma,\tau})\lambda_{\sigma,\tau}{}^{s_{\sigma,\tau}}$$

$$=\text{u.p.}\prod_{\sigma,\tau}(\sigma\mathrm{B}_\tau,\mathfrak{A}_{\sigma,\tau})(\mathrm{B}_{\sigma\tau},\mathfrak{A}_{\sigma,\tau})^{-1}(\mathrm{B}_\sigma,\mathfrak{A}_{\sigma,\tau})(\mathrm{B}_\tau,\mathfrak{P}^{(\sigma^{-1}-1)s_{\sigma,\tau}}).$$

In view of the product over σ and τ we can change the subscripts so that all symbols have B_τ in the left hand argument. Doing this, combining the right hand sides and using (25′) to carry out the product over σ we obtain simply u.p. $\prod_{\tau}(\mathrm{B}_\tau,(A_\tau))$. Since $\partial\mathrm{B}_\tau=\mathfrak{O}^\tau-\mathfrak{O}$ we have then by (18),

$$\text{u.p.}\prod_{\tau}A_\tau(\mathfrak{O}^\tau)A_\tau{}^{-1}(\mathfrak{O})=\text{u.p.}\prod_{\tau}A_\tau{}^{\tau^{-1}-1}(\mathfrak{O})=\prod_{\tau}A_\tau{}^{\tau^{-1}-1}$$

because the function $\prod_{\tau}A_\tau{}^{\tau^{-1}-1}$ is a constant of absolute value 1 (even a root of unity), and is therefore identically equal to the unimodular part of its value at the base point $\mathfrak{O}$. This concludes the proof that $\chi_{\sigma,\tau}$ is a fundamental 2-cocycle.

Now let K be an arbitrary finite extension of k, not necessarily normal. We wish to discuss the relations between characters, cycles, etc. of k, and those of K. For this purpose, we will use lower case letters to denote objects associated with k, the corresponding capitals denoting corresponding objects in K. The surface $R(K)$ is a finite unramified covering of $R(k)$; let π be the projection map of $R(K)$ on $R(k)$. For any point $\mathfrak{Q}\,\varepsilon\,R(K)$, $\pi\mathfrak{Q}$ is the point on $R(k)$ below $\mathfrak{Q}$, and is the norm of $\mathfrak{Q}$ in the sense of divisors. π maps paths B on $R(K)$ into paths $\pi\mathrm{B}$ on $R(k)$. Applied to closed paths through a fixed base point $\mathfrak{O}$ on $R(K)$, this map induces a monomorphism of the fundamental group $F(K)$ of $R(K)$ (based on $\mathfrak{O}$) into the fundamental group $F(k)$ of $R(k)$ (based on $\mathfrak{o}=\pi\mathfrak{O}$). K/k is normal if and ony if $\pi F(K)$ is a normal subgroup of $F(k)$, in which case $G(K/k)\approx F(k)/\pi F(K)$.

There is also a map of 1-cells on the ground surface $R(k)$ into 1-chains on the covering surface $R(K)$, which we call the transfer map and denote by V. If γ is a 1-cell on $R(k)$, then V_γ is the sum of all the 1-cells on $R(K)$ which lie over γ. More precisely, if γ has beginning point $\mathfrak{o}$, and if $\mathfrak{O}_j$, $1 \leqq j \leqq n = [K:k]$, are the points above $\mathfrak{o}$, then for each j there is a unique path Γ_j on $R(K)$ which begins at $\mathfrak{O}_j$ and cover γ; we put $V_\gamma = \sum_{j=1}^{n} \Gamma_j$. Obviously the existence of this map V depends on the finiteness of the covering. V carries cycles into cycles, and induces a homomorphism $V: H_1(k) \rightarrow H_1(K)$. If we identitfy H_1 with the factor commutator group F/F', and view $F(K)$ as subgroup of $F(k)$ by the imbedding π, then it is easy to see that this homomorphism V is none other than the group theoretical transfer or *Verlagerung* from $F(k)$ into the subgroup $F(K)$.

We define the trace map S of differentials of K into those of k, and the cotrace map S^* in the opposite direction by the formulas

$$(27) \qquad (SdU)(\mathfrak{q}) = \sum_{\mathfrak{Q}|\mathfrak{q}} dU(\mathfrak{Q}), \qquad (S^*du)(\mathfrak{Q}) = du(\pi\mathfrak{Q}).$$

From these definitions, the rules

$$(28) \qquad \int_\gamma SdU = \int_{V_\gamma} dU, \qquad \int_\Gamma S^*du = \int_{\pi\Gamma} du,$$

are evident. Concerning our special differentials of the third kind we have

$$(29) \qquad Sd\log\mathfrak{A} = d\log N\mathfrak{A}. \qquad S^*d\log\mathfrak{a} = d\log i\mathfrak{a},$$

where N is the norm map of ideals from K to k and i is the injection map of ideals of k into K. To prove (29) one has simply to observe that the left hand sides have the characterizing properties of the right. Indeed, from (28) we see that the left hand sides have pure imaginary periods, and from the definition of S and S^* it is easily checked that they have the correct residues. Combining (28) and (29) we obtain the following rules for our symbol:

$$(30) \qquad (\beta, N\mathfrak{A})_k = (V\beta, \mathfrak{A})_K; \qquad (\mathrm{B}, i\mathfrak{a})_K = (\pi\mathrm{B}, \mathfrak{a})_k.$$

Choosing a point $\mathfrak{P} \varepsilon R(K)$ with which to describe the characters $\chi_K(\Gamma,\lambda)$ of K, and using $\mathfrak{p} = \pi\mathfrak{P}$ to describe the characters $\chi_k(\gamma,\lambda)$ of k, we find the following formulas for the conorm and norm of characters:

$$(31) \qquad N^*(\chi_k(\gamma,\lambda)) = \chi_K(V_\gamma,\lambda),$$

$$(32) \qquad N(\chi_K(\Gamma,\lambda)) = \chi_k(\pi\Gamma, \lambda^n(\Gamma, \mathfrak{p}\mathfrak{P}^{-n})_K),$$

where $n = [K:k]$. Indeed, from (30), it follows that the left sides give the same values as the right sides when applied to divisors of degree 0, and one can easily check directly that they give the same when applid to $\mathfrak{P}$, resp. $\mathfrak{p}$. Since the map $\lambda \rightarrow \lambda^n$ maps T onto T, (32) shows that the character norm group $N(\hat{E}(K))$ is equal to $\chi_k(\pi H_1(K), T)$, and consequently,

$$\hat{E}(k)/N\hat{E}(K) \approx H_1(k)/\pi H_1(K). \tag{33}$$

If K/k is abelian with group G, then the correspondence which attaches to each element of $H_1(k)$ the corresponding covering transformation of $R(K)$ induces an isomorphism

$$H_1(k)/\pi H_1(K) \approx G. \tag{34}$$

(Covering transformations can be associated with homology classes rather than homotopy classes because the covering is assumed abelian.) We contend that the reciprocity law isomorphism of our pseudo class field theory is obtained by combining the isomorphisms (33) and (34). To prove this, we let $\chi_{\sigma,\tau} = \chi_K(\Gamma_{\sigma,\tau}, \lambda_{\sigma,\tau})$ be a fundamental 2-cocycle for K/k of the type described a few paragraphs above, and we compute the Nakayama map associated with χ. This is the map

$$\rho \rightarrow \prod_\sigma \chi_{\sigma,\rho} = \prod_\sigma \chi_K(\Gamma_{\sigma,\rho}, \lambda_{\sigma,\rho}) = \chi_K(\sum_\sigma \Gamma_{\sigma,\rho}, \prod_\sigma \lambda_{\sigma,\rho}).$$

To work this out we use the definition of $\Gamma_{\sigma,\rho}$ and find

$$\sum_\sigma \Gamma_{\sigma,\rho} = \sum_\sigma (\mathrm{B}_\sigma - \mathrm{B}_{\sigma\rho} + \sigma\mathrm{B}_\rho) = \sum_\sigma \sigma\mathrm{B}_\rho = V\beta_\rho,$$

where $\beta_\rho = \pi\mathrm{B}_\rho$, the paths $\sigma\mathrm{B}_\rho$, $\sigma \varepsilon G$, being then just those paths on $R(K)$ above β_ρ. Combining these last two computations and using (31) we find

$$\prod_\sigma \chi_{\sigma,\rho} = \chi_K(V\beta_\rho, \prod_\sigma \lambda_{\sigma,\rho}) = N^*\chi_k(\beta_\rho, \lambda_\rho),$$

where $\lambda_\rho = \prod_\sigma \lambda_{\sigma,\rho}$. According to the explicit formula (I, 22) of the first half of this paper, we conclude that the reciprocity law attaches the automorphism ρ to the coset of $\chi_k(\beta_\rho, \lambda_\rho) \bmod N\hat{E}(K)$. To prove our contention therefore, we have only to show that ρ is the covering transformation attached to β_ρ. But that is obvious because $\beta_\rho = \pi\mathrm{B}_\rho$, and B_ρ was a path on $R(K)$ leading from the base point $\mathfrak{O}$ to $\mathfrak{O}^\rho$.

Weil [8] has remarked that the kernel of this infinite reciprocity law map of ordinary class field theory is the connected component of the idèle class group. The same is true for our pseudo class field theory. Viewing

$H_1(k)$ as an everywhere dense subgroup of the Galois group $G(A_k/k)$ of the (infinite) maximal abelian extension A_k of k, we see from what we have just proved, that the infinite reciprocity law map is just the natural map $\hat{E}(k) \rightarrow H_1(k)$. Its kernel is isomorphic to T, and is evidently the connected component of $\hat{E}$.

INSTITUTE FOR ADVANCED STUDY,
PRINCETON UNIVERSITY.

REFERENCES.

[1] H. Cartan and S. Eilenberg, *Homological Algebra*, Princeton University Press (forthcoming).

[2] H. Weyl, *Die Idee der Riemannschen Fläche*, Teubner, Leipzig, 1923.

[3] J. Igusa, "Zur klassischen Theorie der algebraischen Functionen," *Journal of the Mathematical Society of Japan*, vol. 1 (1943), pp. 63-72.

[4] E. Artin, *Algebraic Numbers and Alegbraic Functions I*, Mimeographed Notes, New York University, 1951.

[5] G. Hochschild and T. Nakayama, "Cohomology in class field theory," *Annals of Mathematics*, vol. 55 (1952), pp. 348-366.

[6] J. Tate, "The higher dimensional cohomology groups of class field theory," *ibid.*, vol. 56 (1952), pp. 294-297.

[7] E. Artin and J. Tate, *Algebraic Numbers and Algebraic Functions II*, Mimeographed Notes, New York University (in preparation).

[8] A. Weil, "Sur la théorie du corps de classes," *Journal of the Mathematical Society of Japan*, vol. 3 (1951), pp. 1-35.

[9] C. Chevalley, *Class Field Theory*, Nagoya University, Japan, 1954.

LIE GROUPS AND LIE HYPERALGEBRAS OVER A FIELD OF CHARACTERISTIC $p > 0$ (II).*

By JEAN DIEUDONNÉ.

1. Introduction. This paper is a continuation of two earlier ones ([3], [5]); we keep the same notations and terminology. We first complete the general theory outlined in these papers by giving characterizations of a Lie hyperalgebra and of a derived homomorphism; unfortunately, these characterizations do not seem at present to provide very useful tools for further development of the theory. In the remainder of the paper, we first study the formal Lie groups of *dimension one* over an algebraically closed field; we are able to characterize, among those groups, those which are isomorphic to the multiplicative group W_1^*, by the consideration of its Lie *algebra* only; moreover, we can show that the hyperalgebras of all *abelian* Lie groups of dimension one can be reduced to well determined types, forming a denumerable sequence; but the question remains open of the existence of groups corresponding to each of these hyperalgebras.[1] Finally, the same methods yield some information on general abelian formal Lie groups, but (even when the base field is algebraically closed) the complete description of these groups still seems a problem of considerable difficulty.

2. General results. Let G be a formal Lie group of dimension n over a field K of characteristic $p > 0$. In the "Taylor formula" for G ([3], formulas (25) and (28)),

$$f(xy) = \sum_{\alpha} P_\alpha(y) X_\alpha f \tag{1}$$

---

* Received November 6, 1954.

[1] Since this paper was written, M. Lazard has proved that every one-dimensional formal Lie group is necessarily abelian (C. R. Acad. Sci. Paris, vol. 239 (1954), pp. 942-945). On the other hand, I have proved the existence of all one-dimensional abelian Lie groups whose hyperalgebras are described in Theorem 2 (no. 8); the proof will be published in a forthcoming paper in *Mathematische Zeitschrift.*

Added in proof. In a paper to appear in *Bull. Soc. Math. France,* M. Lazard has independently obtained the classification and existence of one-dimensional formal Lie groups over an algebraically closed field, by an entirely different method.

we can rearrange the series in the right-hand member by grouping the terms containing the same monomial y_α in $\mathbf{y}$, and write[2]

$$(2) \qquad f(\mathbf{xy}) = \sum_\alpha y_\alpha Z_\alpha f.$$

From the properties of the series P_α ([3], no. 10), it follows that Z_α is a finite combination of invariant differential operators X_λ of height $h(\lambda) \leqq h(\alpha)$, which can also be characterized as those satisfying the "initial conditions" $Z_\alpha(\mathbf{e}) = (1/\alpha!) D_\alpha(\mathbf{e})$ for every α, and therefore are obviously linearly independent over the ring $\mathfrak{o}$ of formal power series. The arguments used in [3, no. 11] can then be repeated *verbatim* for x_α and Z_α instead of $P_\alpha(\mathbf{x})$ and X_α; in particular, if we write

$$(3) \qquad Z_\alpha Z_\beta = \sum_\gamma d_{\alpha\beta\gamma} Z_\gamma$$

(with the convention $Z_0 = I$, hence $d_{\alpha 0\beta} = d_{0\alpha\beta} = \delta_{\alpha\beta}$ and $d_{\alpha\beta 0} = 0$ if $\alpha \neq 0$ or $\beta \neq 0$), we have

$$(4) \qquad Z_\beta x_\alpha = \sum_\gamma d_{\gamma\beta\alpha} x_\gamma$$

and moreover, the group law is expressed by the series

$$(5) \qquad \phi_i(\mathbf{x}, \mathbf{y}) = \sum_{\alpha,\beta} d_{\alpha\beta\epsilon_i} x_\alpha y_\beta \qquad (1 \leqq i \leqq n)$$

where $\epsilon_i = (0, \cdots, 0, 1, 0, \cdots, 0)$, with 1 at the i-th place.

In addition to these relations, we note that if $h = fg$ is a product of two power series, $h(\mathbf{xy}) = f(\mathbf{xy}) g(\mathbf{xy})$. Let us (partially) order the set of indices $\alpha = (\alpha_1, \cdots, \alpha_n)$ in the usual way, writing $(\alpha_1, \cdots, \alpha_n) \leqq (\beta_1, \cdots, \beta_n)$ to mean $\alpha_i \leqq \beta_i$ for each i. Then, the preceding relation and the "Taylor formula" (2) yield immediately for the operators Z_α, the "generalized Leibniz formula"

$$(6) \qquad Z_\alpha(fg) = \sum_{0 \leqq \beta \leqq \alpha} (Z_\beta f)(Z_{\alpha-\beta} g).$$

Conversely:

PROPOSITION 1. *For each α, let Z_α be a K-linear endomorphism of $\mathfrak{o}$ such that $Z_0 = I$ and that each Z_α is a special semi-derivation of height $h(\alpha)+1$; and suppose these operators verify the relations* (3), (4) *and* (6), *with $d_{\alpha 0\beta} = d_{0\alpha\beta} = \delta_{\alpha\beta}$. Then formulae* (5) *define a group law, for which* (2) *is the Taylor formula.*

From the assumption that the Z_α are semi-derivations, it follows that if $f(\mathbf{x}) = \sum_\lambda a_\lambda x_\lambda$, we can write $Z_\alpha f(\mathbf{x}) = \sum_\lambda a_\lambda Z_\alpha x_\lambda$, by expressing f as a polynomial with coefficients in an $\mathfrak{o}_r$, with $r > h(\alpha)$. Let us define the operator R_y on $\mathfrak{o}$ by the equation

---

[2] The idea to work systematically with the Z_α instead of the X_α occurred to me during several stimulating talks I had with J. Delsarte on this and related topics.

$$R_y f(\boldsymbol{x}) = \sum_{\alpha} y_\alpha Z_\alpha f(\boldsymbol{x}). \tag{7}$$

It follows from (6) that $R_y(fg) = (R_y f)(R_y g)$; on the other hand, one can write $R_y f(\boldsymbol{x}) = \sum_{\lambda} a_\lambda R_y(x_\lambda)$; but

$$R_y(x_\lambda) = (R_y(x_1))^{\lambda_1} \cdots (R_y(x_n))^{\lambda_n}$$

and from (4) it follows that $\phi_i(\boldsymbol{x}, \boldsymbol{y}) = R_y(x_i)$ is given by formula (5) $(1 \leqq i \leqq n)$. We can therefore write

$$R_y f(\boldsymbol{x}) = f(\phi_1(\boldsymbol{x}, \boldsymbol{y}), \cdots, \phi_n(\boldsymbol{x}, \boldsymbol{y})). \tag{8}$$

It follows from (7) that $R_e f(\boldsymbol{x}) = f(\boldsymbol{x})$; on the other hand, the relations (4) and $d_{0\alpha\beta} = \delta_{\alpha\beta}$ show that the value of $Z_\beta x_\alpha$ for $\boldsymbol{x} = \boldsymbol{e}$ is $\delta_{\alpha\beta}$; hence $Z_\alpha f(\boldsymbol{e}) = a_\alpha$, and therefore $R_y f(\boldsymbol{e}) = f(\boldsymbol{y})$; in particular $\phi_i(\boldsymbol{e}, \boldsymbol{x}) = \phi_i(\boldsymbol{x}, \boldsymbol{e}) = x_i$. On the other hand, let us write $L_y f(\boldsymbol{x}) = R_x f(\boldsymbol{y})$; this is again a linear operator on $\mathfrak{o}$, and the associativity conditions are equivalent to the commutativity relation

$$R_y L_z = L_z R_y. \tag{9}$$

To prove (9), we notice that

$$\begin{aligned} R_y L_z f(\boldsymbol{x}) &= \sum_{\alpha, \beta} y_\alpha (Z_\alpha x_\beta)(Z_\beta f(\boldsymbol{z})) \\ &= \sum_{\alpha, \beta, \gamma} d_{\gamma\alpha\beta} x_\gamma y_\alpha (Z_\beta f(\boldsymbol{z})) \end{aligned}$$

due to (4); on the other hand, (3) shows that

$$\begin{aligned} L_z R_y f(\boldsymbol{x}) &= \sum_{\alpha, \beta} x_\alpha y_\beta (Z_\alpha Z_\beta f(\boldsymbol{z})) \\ &= \sum_{\alpha, \beta, \gamma} d_{\alpha\beta\gamma} x_\alpha y_\beta (Z_\gamma f(\boldsymbol{z})) \end{aligned}$$

whence (9), and this ends our proof.

It is easy to verify that the definition (4) of the linear operators Z_α, together with the associativity conditions

$$\sum_{\lambda} d_{\alpha\beta\lambda} d_{\lambda\gamma\delta} = \sum_{\lambda} d_{\beta\gamma\lambda} d_{\alpha\lambda\delta} \tag{10}$$

for the coefficients, imply the "multiplication table" (3). On the other hand, relations (6) are equivalent to the relations obtained by applying them to $f = x_\lambda$, $g = x_\mu$, that is

$$d_{\alpha, \beta, \lambda+\mu} = \sum_{\substack{0 \leqq \gamma \leqq \alpha \\ 0 \leqq \delta \leqq \beta}} d_{\gamma\delta\lambda} d_{\alpha-\gamma, \beta-\delta, \mu} \tag{11}$$

for all indices $\alpha, \beta, \lambda, \mu$. A Lie hyperalgebra may thus be said to consist in a family $(d_{\alpha\beta\gamma})$ of elements of K satisfying (10) and (11), and such that $d_{0\alpha\beta} = d_{\alpha 0\beta} = \delta_{\alpha\beta}$, $d_{\alpha\beta 0} = 0$ except for $\alpha = \beta = 0$, $d_{000} = 1$, and $d_{\alpha\beta\gamma} = 0$ if $x_\gamma \varepsilon \mathfrak{o}_r$, and $r > h(\gamma)$; but it is obvious that such a description is hardly a workable one.

3. Let now $\boldsymbol{u}$ be a homomorphism of G into a formal Lie group $\bar{G}$ of dimension m; the arguments of [3, no. 12] can be applied to the x_α and Z_α and show that one can write

$$\boldsymbol{u}'(Z_\alpha) = \sum_\lambda a_{\alpha\lambda} \bar{Z}_\lambda \tag{12}$$

and

$$\bar{f}(\boldsymbol{u}(\boldsymbol{x})) = \sum_\alpha x_\alpha (\boldsymbol{u}'(Z_\alpha)(\bar{\boldsymbol{e}}))\bar{f} = \sum_\lambda (\boldsymbol{u}(\boldsymbol{x}))_\lambda \bar{Z}_\lambda (\bar{\boldsymbol{e}})\bar{f} \tag{13}$$

where $(\boldsymbol{u}(\boldsymbol{x}))_\lambda$ means $\bar{x}_\lambda$ in which $u_i(\boldsymbol{x})$ has been substituted for $\bar{x}_i$ $(1 \leqq i \leqq m)$. In addition, if $\bar{h} = \bar{f}\bar{g}$, one has $\bar{h}(\boldsymbol{u}(\boldsymbol{x})) = \bar{f}(\boldsymbol{u}(\boldsymbol{x}))\bar{g}(\boldsymbol{u}(\boldsymbol{x}))$ and therefore formula (13) yields the conditions

$$\boldsymbol{u}'(Z_\alpha)(\bar{f}\bar{g}) = \textstyle\sum_{0 \leqq \beta \leqq \alpha} (\boldsymbol{u}'(Z_\beta)(\bar{f}))(\boldsymbol{u}'(Z_{\alpha-\beta})(\bar{g})). \tag{14}$$

Conversely:

Proposition 2. *Let $\boldsymbol{u}'$ be a homomorphism of the Lie hyperalgebra of G into that of $\bar{G}$ satisfying conditions* (14), *transforming the unit element into the unit element and sending each $\mathfrak{g}_r$ into $\bar{\mathfrak{g}}_r$. Then $\boldsymbol{u}'$ is the derived homomorphism of a homomorphism $\boldsymbol{u}$ of G into $\bar{G}$.*

Let us define a mapping C of $\mathfrak{o}(\bar{G})$ into $\mathfrak{o}(G)$ by the formula

$$C\bar{f} = \sum_\alpha x_\alpha (\boldsymbol{u}'(Z_\alpha)(\bar{\boldsymbol{e}}))\bar{f}; \tag{15}$$

it follows from (14) that $C(\bar{f}\bar{g}) = C(\bar{f})C(\bar{g})$, in other words C is a ring homomorphism. If $\boldsymbol{u}'(Z_\alpha)$ is given by (12), let us define

$$u_i(\boldsymbol{x}) = C(\bar{x}_i) = \sum_\alpha a_{\alpha\epsilon_i} x_\alpha \quad \text{for } 1 \leqq i \leqq m. \tag{16}$$

Then, from (15) and (12) it follows that

$$C(\bar{x}_\lambda) = (u_1(\boldsymbol{x}))^{\lambda_1} \cdots (u_m(\boldsymbol{x}))^{\lambda_m} = \sum_\alpha a_{\alpha\lambda} x_\alpha. \tag{17}$$

Formula (15) and the assumption that $\boldsymbol{u}'$ sends $\mathfrak{g}_r$ into $\bar{\mathfrak{g}}_r$ for each r imply that C transforms $\mathfrak{o}_r(\bar{G})$ into a subring of $\mathfrak{o}_r(G)$. If we write the Taylor formula in $\bar{G}$

$$\bar{f}(\bar{\boldsymbol{x}}) = \sum_\lambda \bar{x}_\lambda \bar{Z}_\lambda (\bar{\boldsymbol{e}})\bar{f}$$

we deduce from the preceding remarks that one has

$$C\bar{f}=\sum_{\lambda} C(\bar{x}_{\lambda})\bar{Z}_{\lambda}(\bar{\boldsymbol{e}})\bar{f}=\bar{f}(\boldsymbol{u}(\boldsymbol{x})).$$

To prove that $\boldsymbol{u}$ is a homomorphism of G into $\bar{G}$, we have to prove the relation

$$CR_{\boldsymbol{u}(\boldsymbol{y})}=R_{\boldsymbol{y}}C. \tag{18}$$

But one has

$$R_{\boldsymbol{u}(\boldsymbol{y})}\bar{f}=\sum_{\lambda,\mu}\bar{x}_{\lambda}(\boldsymbol{u}(\boldsymbol{y}))_{\mu}(\bar{Z}_{\lambda}\bar{Z}_{\mu}(\bar{\boldsymbol{e}}))\bar{f}$$

hence

$$CR_{\boldsymbol{u}(\boldsymbol{y})}\bar{f}=\sum_{\lambda,\mu}(\boldsymbol{u}(\boldsymbol{x}))_{\lambda}(\boldsymbol{u}(\boldsymbol{y}))_{\mu}(\bar{Z}_{\lambda}\bar{Z}_{\mu}(\bar{\boldsymbol{e}}))\bar{f}.$$

Therefore, (17) and the assumption that $\boldsymbol{u}'$ is a homomorphism enable one to write

$$\begin{aligned}CR_{\boldsymbol{u}(\boldsymbol{y})}\bar{f}&=\sum_{\alpha,\beta}x_{\alpha}y_{\beta}(\boldsymbol{u}'(Z_{\alpha}Z_{\beta})(\bar{\boldsymbol{e}}))\bar{f}\\&=\sum_{\alpha,\beta,\gamma}d_{\alpha\beta\gamma}x_{\alpha}y_{\beta}(\boldsymbol{u}'(Z_{\gamma})(\bar{\boldsymbol{e}}))\bar{f}\end{aligned}$$

and, using (4)

$$CR_{\boldsymbol{u}(\boldsymbol{y})}\bar{f}=\sum_{\beta,\gamma}y_{\beta}(Z_{\beta}x_{\gamma})(\boldsymbol{u}'(Z_{\gamma})(\bar{\boldsymbol{e}}))\bar{f};$$

but, from (15), this is also

$$\sum_{\beta}y_{\beta}(Z_{\beta}C\bar{f})=R_{\boldsymbol{y}}C\bar{f}$$

using formula (7), and this ends the proof.

4. In what follows, we will use repeatedly the notion of *canonical group law*; we will give here a brief summary of the definitions and results pertaining to this notion (for the proofs, see [5]). To say that the group law of G is canonical means that the power series $P_{0i}(\boldsymbol{x})=x_i$ for $1\leqq i\leqq n$; from (2), it follows that this is equivalent to the following property of the Z_{α}: *if* α *is distinct from* ϵ_i $(1\leqq i\leqq n)$, *the coefficient of each* X_{0j} $(1\leqq j\leqq n)$ *in the expression of* Z_{α} *as linear combination of the* X_{λ} *is* 0.

When the group law of G is not canonical, there exists an isomorphism $\boldsymbol{u}$ of G onto a group $\bar{G}$ having canonical law. Such an isomorphism can be obtained by a "standard process" described in detail in [5]: one considers an infinite sequence $(G^{(r)})$ of groups, with $G^{(0)}=G$, and for each r one defines an isomorphism $\boldsymbol{u}^{(r)}$ of $G^{(r)}$ onto $G^{(r+1)}$ by the formulae $u_i^{(r)}(\boldsymbol{x})=P_{0i}^{(r)}(\boldsymbol{x})$ for $1\leqq i\leqq n$; it turns out that $u_i^{(r)}(\boldsymbol{x})=x_i+v_i^{(r)}(\boldsymbol{x})$ where the monomials in $v_i^{(r)}$ have all a height $\geqq r$, and moreover there are no monomials of type $x_j^{p^k}$ for $k\geqq 1$ (see [3, formula (34)]). The group $\bar{G}$ with canonical law and the isomorphism $\boldsymbol{u}$ are obtained by a "passage to the limit" from $G^{(r)}$ and the composite isomorphism $\boldsymbol{u}^{(r)}\boldsymbol{u}^{(r-1)}\cdots\boldsymbol{u}^{(0)}$.

The importance of the notion of canonical law lies chiefly in the *uniqueness theorem* [5, Th. 2] which can be stated in the slightly more general following form (following from the proof of Theorem 2 in [5]): *if the group laws of two n-dimensional formal Lie groups G, H are canonical, and if the constants* $c_{\alpha\beta\gamma}$ *which determine the multiplication of the* X_α *in the hyperalgebras of G and H* (see [3, no. 11]) *are the same in both groups for* $h(\alpha) < r$, $h(\beta) < r$ (and necessarily $h(\gamma) < r$) *then the constants* $d_{\alpha\beta\gamma}$ (see formula (3)) *are also the same in G and H for* $h(\alpha) < r$ *and* $h(\beta) < r$.

Finally, we prove the following

PROPOSITION 3. *If the group law is canonical, one has* $Z_\alpha = (1/\alpha!)X_\alpha$ *for all indices* α *of height* $h(\alpha) = 0$.

We will use induction on the total degree $|\alpha|$, the proposition being obvious for $|\alpha| = 1$. Let $\alpha = (0,\cdots,0,\alpha_i,\alpha_{i+1},\cdots,\alpha_n)$, α_i being the first component $\neq 0$; we can therefore write $\alpha = \alpha_i\epsilon_i + \beta$, where $\beta = (0,\cdots,0,\alpha_{i+1},\cdots,\alpha_n)$. The proposition will be proved if we establish that the difference $Z_\alpha - \frac{1}{\alpha!}X_\alpha$ is a *derivation*; for it will then have to be a linear combination of the X_{0j} $(1 \leqq j \leqq n)$, and if $|\alpha| > 1$, this linear combination must reduce to 0, due to the assumption that the group law is canonical.

Now we have, from (6) and the induction hypothesis

$$Z_\alpha(fg) = f\cdot Z_\alpha g + g\cdot Z_\alpha f + \sum_{(\lambda_i,\gamma)} \Big(\frac{1}{\lambda_i!}X_{0i}^{\lambda_i}Z_\gamma f\Big)\Big(\frac{1}{(\alpha_i-\lambda_i)!}X_{0i}^{\alpha_i-\lambda_i}Z_{\beta-\gamma}g\Big)$$

where in the summation, the pair (λ_i,γ) ranges over the set satisfying the conditions $0 \leqq \lambda_i \leqq \alpha_i$ and $0 \leqq \gamma \leqq \beta$, with the exception of the two pairs $(0,0)$ and (α_i,β). On the other hand, we have similarly

$$\frac{1}{\alpha!}X_\alpha = \frac{1}{\alpha_i}X_{0i}Z_{\alpha-\epsilon_i}$$

and

$$Z_{\alpha-\epsilon_i}(fg) = \sum_{(\mu_i,\gamma)} \Big(\frac{1}{\mu_i!}X_{0i}^{\mu_i}Z_\gamma f\Big)\Big(\frac{1}{(\alpha_i-\mu_i-1)!}X_{0i}^{\alpha_i-\mu_i-1}Z_{\beta-\gamma}g\Big)$$

where in the summation, $0 \leqq \mu_i \leqq \alpha_i - 1$, $0 \leqq \gamma \leqq \beta$; therefore

$$\frac{1}{\alpha!}X_\alpha(fg) = \frac{1}{\alpha_i}\sum_{(\mu_i,\gamma)} \Big(\Big(\frac{1}{\mu_i!}X_{0i}^{\mu_i+1}Z_\gamma f\Big)\Big(\frac{1}{(\alpha_i-\mu_i-1)!}X_{0i}^{\alpha_i-\mu_i-1}Z_{\beta-\gamma}g\Big)$$
$$+ \Big(\frac{1}{\mu_i!}X_{0i}^{\mu_i}Z_\gamma f\Big)\Big(\frac{1}{(\alpha_i-\mu_i-1)!}X_{0i}^{\alpha_i-\mu_i}Z_{\beta-\gamma}g\Big)\Big)$$

with the same range in the summation. As

$$\frac{1}{\alpha_i}\left(\frac{1}{\mu_i!(\alpha_i-\mu_i-1)!}+\frac{1}{(\mu_i-1)!(\alpha_i-\mu_i)!}\right)=\frac{1}{\mu_i!(\alpha_i-\mu_i)!}$$

we have

$$Z_\alpha(fg)-\frac{1}{\alpha!}X_\alpha(fg)=f(Z_\alpha g-\frac{1}{\alpha!}X_\alpha g)+g(Z_\alpha f-\frac{1}{\alpha!}X_\alpha f)$$

and the proof is complete.

This result corresponds of course to the "exponential" expression of the Taylor formula in canonical coordinates, in the classical case; but examples show that there is no such "universal" expression of Z_α as a polynomial in the X_λ, which would be valid for *every* canonical group law, when $h(\alpha)>1$.

5. One-dimensional groups. We first prove two lemmas which will be used repeatedly later on.

LEMMA 1. *Suppose Z_λ and Z_μ are permutable for the pairs (λ,μ) such that, either $0\leqq\lambda\leqq\alpha$, $0\leqq\mu<\beta$, or $0\leqq\lambda<\alpha$, $0\leqq\mu\leqq\beta$; then $[Z_\alpha,Z_\beta]$ is a derivation.*

Indeed, it follows from (6) that

$$Z_\alpha Z_\beta(fg)=\sum_{\substack{0\leqq\lambda\leqq\alpha\\ 0\leqq\mu\leqq\beta}}(Z_\lambda Z_\mu f)(Z_{\alpha-\lambda}Z_{\beta-\mu}g)$$

and

$$Z_\beta Z_\alpha(fg)=\sum_{\substack{0\leqq\lambda\leqq\alpha\\ 0\leqq\mu\leqq\beta}}(Z_\mu Z_\lambda f)(Z_{\beta-\mu}Z_{\alpha-\lambda}g)$$

But, in the range of summation, one has $Z_\lambda Z_\mu=Z_\mu Z_\lambda$ except for the pair $\lambda=\alpha$, $\mu=\beta$; hence

$$(Z_\alpha Z_\beta-Z_\beta Z_\alpha)(fg)=f\cdot(Z_\alpha Z_\beta-Z_\beta Z_\alpha)g+g\cdot(Z_\alpha Z_\beta-Z_\beta Z_\alpha)f$$

which ends the proof.

LEMMA 2. *Suppose the Z_λ corresponding to the range $0\leqq\lambda\leqq\alpha$ are mutually permutable; then*

$$Z_\alpha{}^p(fg)=\sum_{0\leqq\lambda\leqq\alpha}(Z_\lambda{}^p f)(Z_{\alpha-\lambda}{}^p g). \tag{19}$$

It will be enough to prove that, for every integer m, one has

$$Z_\alpha{}^m(fg)=\left(\sum_{0\leqq\lambda\leqq\alpha}(Z_\lambda f)(Z_{\alpha-\lambda}g)\right)^{(m)}$$

where, in the "symbolic power", a product $(Z_{\lambda_1}f)(Z_{\lambda_2}f)\cdots(Z_{\lambda_k}f)$ has to be replaced by $(Z_{\lambda_1}Z_{\lambda_2}\cdots Z_{\lambda_k})f$, and similarly for the products of $Z_\mu g$; by induction on m, this follows immediately from (6) and the assumption that all the Z_λ which intervene commute.

6. From now on until **12**, we shall suppose that the field K is *algebraically closed.* For a one-dimensional group G, we shall write X_h instead of X_{h1}; we have therefore $X_0{}^p = aX_0$; if $a \neq 0$, the change of variable $\bar{x}_1 = u(x_1) = \lambda x_1$ is such that $u'(X_0) = \lambda \bar{X}_0$, hence $u'(X_0{}^p) = \lambda^p \bar{X}_0{}^p = a\lambda \bar{X}_0$. It is therefore possible to choose $\lambda \varepsilon K$ such that one is reduced to the case $a = 1$.

Case A. $X_0{}^p = X_0$. We may suppose [5] that the group law of G is *canonical.* Our first goal is to prove that X_1 and X_0 commute; from Prop. 3 it follows that Lemma 1 can be applied to the indices 1 and p, and therefore

$$X_1 X_0 - X_0 X_1 = cX_0 \tag{20}$$

is a derivation; by induction on k, we get

$$X_1 X_0{}^p - X_0{}^k X_1 X_0{}^{p-k} = kcX_0 \tag{21}$$

and in particular, for $k = p$

$$X_1 X_0{}^p - X_0{}^p X_1 = 0. \tag{22}$$

Comparing (20) and (22) and using the assumption $X_0{}^p = X_0$ yields $c = 0$.

7. We shall now define an inductive process which will ultimately give us the structure of G.

We start from a one-dimensional group $G^{(r)}$, with canonical group law, and we suppose that: 1. $X_0{}^{(r)}, X_1{}^{(r)}, \cdots, X_{r-1}{}^{(r)}$ *are mutually permutable*; 2. $(X_h{}^{(r)})^p = X_h{}^{(r)}$ *for* $0 \leqq h \leqq r-1$. We want to define an isomorphism $v^{(r)}$ of $G^{(r)}$ on a similar group $G^{(r+1)}$, such that $v^{(r)}(x^{(r)}) = x^{(r)} + \mu(x^{(r)})^{p^r} + \cdots$ (terms omitted of degree $> p^r$).

For convenience, we drop the upper index r, and speak therefore of G instead of $G^{(r)}$. We first observe that, as the Z_α for $\alpha < p^r$ are polynomials in $X_0, \cdots, X_{r-1}$, they are mutually permutable. Lemma 1 may therefore be applied, first to X_0 and X_r, which gives

$$X_r X_0 - X_0 X_r = b_0 X_0 \tag{23}$$

and the same argument as in **6** shows that $b_0 = 0$. Suppose we have proved that X_r commutes with $X_0, X_1, \cdots, X_{k-1}$; as the Z_α for $\alpha < p^k$ are polynomials in $X_0, \cdots, X_{k-1}$, they commute with X_r; Lemma 1 may again be applied to X_k and X_r, which gives

$$X_r X_k - X_k X_r = b_k X_0. \tag{24}$$

We treat this equation in the same way as (20), writing first

$$X_r X_k^p - X_k X_r X_k^{p-1} = b_k X_0 X_k^{p-1}.$$

By induction on h, and using the fact that X_0 and X_k commute, we get

$$X_r X_k^p - X_k^h X_r X_k^{p-h} = h b_k X_0 X_k^{p-1} \qquad (1 \leqq h \leqq p) \tag{25}$$

and in particular, for $h = p$, and using the relation $X_k^p = X_k$, we see that X_k and X_r commute. We have thus shown that X_r commutes with X_k for $0 \leqq k \leqq r-1$.

The next step consists first in remarking that $Z_\alpha^p = Z_\alpha$ for $\alpha \leqq p^r - 1$. This is clear by assumption for $\alpha = 1, p, \cdots, p^{r-1}$; it is proved by induction on α for the other values of α. Indeed, we have, by Lemma 2 and the induction hypothesis

$$Z_\alpha^p(fg) = f \cdot Z_\alpha^p g + g \cdot Z_\alpha^p f + \textstyle\sum_{1 \leqq \lambda \leqq \alpha-1} (Z_\lambda f)(Z_{\alpha-\lambda} g) \tag{26}$$

from which one deduces immediately that $Z_\alpha^p - Z_\alpha$ is a *derivation.* But in the expression of Z_α as a linear combination of the X_λ $(\lambda \leqq \alpha)$, there is no term in X_0 (from the assumption that the group law is canonical); as $X_k^p = X_k$ for every $k < r$ (from the inductive assumption), we have necessarily $Z_\alpha^p - Z_\alpha = 0$. This result being obtained, we can apply Lemma 2 to $\alpha = p^r$, since X_r commutes with the Z_α of index $\alpha < p^r$; we obtain

$$X_r^p(fg) = f \cdot X_r^p g + g \cdot X_r^p f + \textstyle\sum_{1 \leqq \lambda \leqq p^r} (Z_\lambda f)(Z_{p^r-\lambda} g)$$

and therefore $X_r^p - X_r$ is a *derivation,* in other words

$$X_r^p - X_r = b X_0. \tag{27}$$

We now consider the transformation $v_1(x) = x + \lambda x^{p^r}$; this is an isomorphism of G on a group G_1, and formulae (12) and (17) show that

$$\begin{cases} v'_1(X_k) = X_k^{(1)} & \text{for} \quad 0 \leqq k \leqq r-1 \\ v'_1(X_r) = X_r^{(1)} + \lambda X_0^{(1)}. \end{cases} \tag{28}$$

Equation (27) yields therefore

$$(X_r^{(1)})^p - X_r^{(1)} = (b + \lambda - \lambda^p) X_0^{(1)}$$

and it is possible to choose λ such that the right-hand side vanishes. The group law of G_1 is not perhaps a canonical one; but from formula (13) and the definition of v_1, it follows at once that for $\alpha < p^r$, $v'_1(Z_\alpha) = Z_\alpha^{(1)}$; formulae (28) prove therefore that the expression of $Z_\alpha^{(1)}$ for $\alpha < p^r$, as linear combination of the $X_\lambda^{(1)}$, has the same coefficients as those in the expression of Z_α as combination of the X_λ. The first $Z_\alpha^{(1)}$ which may contain X_0 are therefore of degree $\alpha > p^r$, in other words, one has

$$P^{(1)}(x^{(1)}) = x^{(1)} + a(x^{(1)})^{p^r+1} + \cdots. \tag{29}$$

The standard method, which yields a canonical group law from an arbitrary one (see **4**), shows then that there exists an isomorphism v_2 of G_1 onto a group $\bar{G}$ with canonical group law, such that $v'_2(X_k^{(1)}) = \bar{X}_k$ for $0 \leqq k \leqq r$, and in addition the isomorphism $v = v_2 v_1$ of G on $\bar{G}$ has the form $v(x) = x + \lambda x^{p^r} + \cdots$.

Returning to our original notations, we have therefore proved the existence of the isomorphism $v^{(r)}$. Starting now from $G = G^{(0)}$, we consider the isomorphism $u^{(r)} = v^{(r)} v^{(r-1)} \cdots v^{(0)}$ of G on $G^{(r+1)}$. It is clear that, in the formal power series $u^{(r)}(x)$, the terms of degree $< p^m$ are *the same* for all $r > m$; we can therefore define a power series u by the condition that all terms of u of degree $< p^m$ are those of every series $u^{(r)}$ for $r > m$. It follows that $w^{(r)} = u(u^{(r)})^{-1}$ is an isomorphism of $G^{(r)}$ on a group $\bar{G}$, such that $w^{(r)}(x^{(r)}) = x^{(r)} + \lambda(x^{(r)})^{p^r} + \cdots$; from which one concludes immediately that the group law of $\bar{G}$ is canonical, that $\bar{G}$ has an abelian hyperalgebra, and therefore is *abelian*, and that $\bar{X}_k^p = \bar{X}_k$ for every integer $k \geqq 0$.

As every formal Lie group over a field K can be considered as a Lie group over any extension of K, we have proved the following

THEOREM 1. *Any formal Lie group G of dimension 1 over a field K such that $X_0^p = \lambda X_0$, with $\lambda \neq 0$, is abelian.*[1] *If in addition K is algebraically closed, G is isomorphic to the multiplicative group W_1^** (with group law $(x, y) \to x + y + xy$); *moreover, there is one and only one group I_0 over K, isomorphic to W_1^*, having a canonical law and whose hyperalgebra is determined by the relations $X_h^p = X_h$ $(h = 0, 1, 2, \cdots)$.*

8. Continuing our study of one-dimensional groups, we turn now to

Case B. $X_0^p = 0$. Here, I am not able to prove that G is abelian.[1] All I can do in that direction is to prove that X_0 *commutes with every other* X_r. Indeed, suppose we have proved that X_0 commutes with $X_1, \cdots, X_{r-1}$; then X_0 commutes with the Z_α such that $\alpha < p^r$, and we can apply Lemma 1 to X_0 and X_r, obtaining

$$X_r X_0 - X_0 X_r = b X_0$$

which can also be written

$$X_r X_0 = X_0 (X_r + b)$$

and from which we deduce, by induction,

$$X_r^k X_0 = X_0 (X_r + b)^k$$

and in particular, for $k = p$

$$X_r^p X_0 - X_0 X_r^p = b^p X_0. \tag{30}$$

Now, one can write $X_r^p = aX_r + Y$, where Y is a linear combination of the X_α such that $\alpha < p^r$. Let σ be the isomorphism $\xi \rightarrow \xi^{p^r}$ of K onto K^{p^r}, and denote also by σ the corresponding mappings of G onto G^σ and of the hyperalgebra of G onto that of G^σ obtained by applying σ to the coefficients of the group law; it is clear that the group law of G^σ is also canonical. Now, if $\boldsymbol{p}'$ is the derived "Frobenius homomorphism [3, no. 13], apply the ring-homomorphism $\sigma^{-1} \circ \boldsymbol{p}'^r$ to the relation $X_r^p = aX_r + Y$; we obtain $X_0^p = a^{-p^r} X_0$, hence $a = 0$, and the inductive assumption shows that X_0 commutes with X_r.

Using Lemma 2, it is then possible to show that $X_1^p = cX_0$, and using Lemma 1, that $[X_1, X_2] = bX_0$; but there does not seem to be any similar argument which will prove that $b = 0$.

We will therefore *assume*, from now on, that G is abelian and has a canonical group law.[1] Suppose we have $X_0^p = X_1^p = \cdots = X_{r-1}^p = 0$; then it is clear that $Z_\alpha^p = 0$ for $\alpha < p^r$, and Lemma 2 shows that X_r^p is a *derivation*, in other words $X_r^p = bX_0$. If $b = 0$, we can go on with the induction; two cases are therefore possible:

Case B1. $X_r^p = 0$ *for every* r. These relations are obviously invariant under any isomorphism, hence we can suppose the group law of G is canonical, and then [3, formula (30)], G is the *additive group* W_1 (with group law $(x, y) \rightarrow x + y$).

Case B2. $X_0^p = X_1^p = \cdots = X_{h-1}^p = 0$, $X_h^p \neq 0$. A transformation of type $\bar{x} = \lambda x$ reduces to the case $X_h^p = X_0$. As a basis for an inductive process, suppose we have also $X_{h+1}^p = X_1, \cdots, X_{h+r-1}^p = X_{r-1}$. Apply the isomorphism $\sigma : \xi \rightarrow \xi^p$ to G. As the group law of G^σ is canonical and the constants of structure of the algebra $\mathfrak{g}_{h+r}$ are in the prime field of characteristic p, and therefore $\mathfrak{g}_{h+r}^\sigma$ is identical to $\mathfrak{g}_{h+r}$, it follows from the uniqueness theorem (see 4), that the group laws of G and G^σ coincide modulo x^{p^r} and y^{p^r}; hence the coefficients, in the group law of G, of monomials $x^\lambda y^\mu$, where both λ and μ

are $< p^r$, belong to the prime field. Now the homomorphism $Y \rightarrow Y^p$ of the (commutative) ring $\mathfrak{S}_{h+r}$ coincides with the homomorphism $\boldsymbol{p}'^h$ on the generators X_i $(0 \leqq i \leqq h + r - 1)$ of the algebra $\mathfrak{S}_{h+r}$; as the Z_α for $\alpha < p^{r+h}$ are polynomials in those X_i with coefficients in the prime field, it follows from the definition of $\boldsymbol{p}'$ and of the Z_α, that we have $Z_\alpha{}^p = 0$ if α is not a multiple of p^h, and $Z_\alpha{}^p = Z_\beta$ if $\alpha = p^h\beta$. Lemma 2 then proves that $\mathrm{X}_{h+r}{}^p - X_r$ is a *derivation*, in other words,

$$X_{h+r}{}^p - X_r = bX_0. \tag{31}$$

We can then consider an isomorphism of type $v_1(x) = x + \lambda x^{p^r}$ on a group G_1, and we have

$$v'_1(X_k) = X_k{}^{(1)} \quad \text{for} \quad 0 \leqq k \leqq r - 1$$

$$v'_1(X_r) = X_r{}^{(1)} + \lambda X_0{}^{(1)}$$

$$v'_1(X_{k+r}) = X_{k+r}{}^{(1)} + \phi_k(X_0{}^{(1)}, \cdots, X_{k-1}{}^{(1)}) + \lambda^{p^k} X_k{}^{(1)}$$

for $1 \leqq k \leqq h$, ϕ_k being a polynomial in the $X_j{}^{(1)}$. This, together with equation (31), gives $(X_{k+r}{}^{(1)})^p = X_{k+r-h}{}^{(1)}$ for $k < h$, and

$$(X_{h+r}{}^{(1)})^p - X_r{}^{(1)} = (b + \lambda - \lambda^{p^{h+1}}) X_0{}^{(1)}$$

and therefore we can choose λ such that the right-hand side vanishes. We can then, as in **7**, determine an isomorphism v_2 of G_1 onto a group $\bar{G}$ with canonical group law, such that $v'_2(X_k{}^{(1)}) = \bar{X}_k$ for $k \leqq h + r$, and in addition the isomorphism $v = v_2 v_1$ of G on $\bar{G}$ has the form $v(x) = x + \lambda x^{p^r} + \cdots$. The argument is then concluded by induction on r, as in **7**, and yields the following result:

THEOREM 2. *Let G be a one-dimensional abelian formal Lie group over an algebraically closed field K, such that $X_0{}^p = 0$. If $X_h{}^p = 0$ for every $h \geqq 0$, G is isomorphic to the additive group W_1. If on the contrary there is a smallest integer $r > 0$ such that $X_r{}^p \neq 0$, there is one and only one group I_r over K, isomorphic to G, having a canonical law, and whose hyperalgebra is determined by the relations*

$$X_k{}^p = 0 \quad \textit{for} \quad k < r, \qquad X_k{}^p = X_{k-r} \quad \textit{for} \quad k \geqq r. \tag{32}$$

9. The elliptic curves. Theorems 1 and 2 would give a complete description of all abelian Lie groups of dimension 1 over an algebraically closed field, *if we knew that for every $r > 0$ there exists a group G such that $X_k{}^p = 0$ for $k < r$ and $X_r{}^p \neq 0$.* But at present I am unable to exhibit such groups,[1] except for the case $r=1$, $p=5$: the computations in [4] show indeed that the elliptic group $E_5(0,b)$, with $b\neq 0$, is of such a type. It seems likely that the groups I_r exist for every $r>0$ and every prime p.

It is possible to give a simple criterion for an elliptic group $E_p(a,b)$ $(p\geqq 3)$ to be isomorphic to the multiplicative group $W_1{}^*$, when K is algebraically closed. More generally, let G be a one-dimensional formal Lie group, with the group law written

$$(x,y)\rightarrow\phi(x,y)=x+v_1(x)y+v_2(x)y^2+\cdots \tag{33}$$

and

$$v_1(x)=1+a_1x+\cdots.$$

The criterion of Theorem 1 can be transformed in the following one:

PROPOSITION 4. *In order that G be isomorphic to the multiplicative group $W_1{}^*$ (K being an algebraically closed field), a necessary and sufficient condition is that the coefficient of x^{p-1} in the power series $1/v_1(x)$ be different from 0.*

Indeed, we have

$$X_0=v_1D_0;$$

it follows that

$$X_0{}^p=((v_1D_0)^{p-1}v_1)\cdot D_0+\cdots$$

where the unwritten terms are combinations of $D_0{}^2, D_0{}^3,\cdots$; as we know that $X_0{}^p$ is a derivation, these terms are in fact reduced to 0, and the constant γ such that $X_0{}^p=\gamma X_0$ is the constant term in the series $(v_1D_0)^{p-1}v_1$. Our assertion will then result of the following lemma:

LEMMA 3. *Let D be a derivation in a commutative algebra A of characteristic $p>0$; then, for every element z in that algebra,*

$$(zD)^{p-1}z=-zD^{p-1}(z^{p-1}). \tag{34}$$

Supposing this proved, and replacing z by v_1 and D by D_0 in (34), we see that γ is the constant term of $-D^{p-1}(v_1{}^{p-1})$, or equivalently, the coefficient of x^{p-1} in $v_1{}^{p-1}$; but as $v_1{}^p(x)=1+a_1{}^px^p+\cdots$, γ is also the coefficient of x^{p-1} in $v_1{}^{-1}$, and this ends the proof of Prop. 4.

We are thus reduced to proving Lemma 3. I am indebted to G. Hochschild for the following proof, which is shorter and more elegant than my original one, and which I reproduce here with his kind permission.

Hochschild's proof is based on the following more general result:

LEMMA 4. *Let R be any commutative algebra over the prime field $\boldsymbol{F}_p$, and let δ be a derivation of R. Then, for any $r \varepsilon R$,*

$$(r\delta)^p = r^p\delta^p + (r\delta)^{p-1}(r)\delta \tag{35}$$

(here $r\delta$ stands for the derivation $x \rightarrow r\delta(x)$).

Let t and x_i $(i=0,1,\cdots)$ be indeterminates, and let H be the algebra of polynomials $\boldsymbol{F}_p[t, x_0, x_1, \cdots]$. There is a derivation τ of H such that $\tau(x_i) = x_{i+1}$, $\tau(t) = 1$. Identifying an element g of H with the multiplication $f \rightarrow gf$ by this element, we have (in the ring of endomorphisms of the vector space H)

$$\tau x_i = x_i\tau + x_{i+1}.$$

These relations show immediately that we can write

$$(x_0\tau)^p = x_0{}^p\tau^p + \sum_{i=1}^{p-1} q_i\tau^i \tag{36}$$

where the q_i are polynomials in the x_j. Applying the endomorphisms on both sides of (36) to the element t^{p-1}, we obtain (observing that $(x_0\tau)^p$ is a derivation)

$$(p-1)(x_0\tau)^p(t)t^{p-2} = \sum_{i=1}^{p-1}(p-1)\cdots(p-i)q_it^{p-1-i}.$$

This implies $q_i = 0$ for $i>1$ and $q_1 = (x_0\tau)^p(t) = (x_0\tau)^{p-1}(x_0)$. Hence

$$(x_0\tau)^p = x_0{}^p\tau^p + (x_0\tau)^{p-1}(x_0). \tag{37}$$

Let S be the subring of the ring of endomorphisms of the vector space H, generated by τ and the multiplications by the x_i. There is a homomorphism f of S into the endomorphism ring of the vector space R, such that $f(\tau) = \delta$, $f(x_0)$ is the multiplication by r, and $f(x_i)$ the multiplication by $\delta^i(r)$ for $i \geqq 1$. Applying f to both sides of (37) yields relation (35).

Lemma 4 being proved, we now apply it to $R = \boldsymbol{F}_p(t, x_0, x_1, \cdots)$ (field of quotients of H), and to the derivation δ of R such that $\delta(x_i) = x_{i+1}$ and $\delta(t) = 1/x_0$. Taking the values of both sides of (35), with $r = x_0$, for the element t, gives

$$0 = x_0{}^p\delta^{p-1}(1/x_0) + (x_0\delta)^{p-1}(x_0)\cdot(1/x_0),$$

whence

$$(x_0\delta)^{p-1}(x_0) = -x_0{}^{p+1}\delta^{p-1}(1/x_0).$$

Since $\delta(x_0{}^p) = 0$, this may be written

$$(x_0\delta)^{p-1}(x_0) = -x_0\delta^{p-1}(x_0{}^{p-1}). \tag{38}$$

Now there is a homomorphism ϕ of the ring of polynomials $F_p[x_0, x_1, \cdots]$ into the algebra A such that $\phi(x_0) = z$, $\phi(x_i) = D^i(z)$ and $\phi \circ \delta = D \circ \phi$. Applying ϕ to both sides of (38) gives (34), and this ends the proof of Lemma 3.

For the elliptic group $E_p(a, b)$ [4], we have $v_1(x) = (1 - ax^4 - bx^6)^{\frac{1}{2}}$; to say that the coefficient of x^{p-1} in $1/v_1(x)$ is $\neq 0$ means that the *Hasse invariant* [6] of the elliptic curve is $\neq 0$.[3] Hence:

COROLLARY. *Let K be an algebraically closed field of characteristic $p > 3$. In order that the elliptic group $E_p(a, b)$ be isomorphic (as a formal Lie group) to the multiplicative group W_1^*, it is necessary and sufficient that the Hasse invariant of $E_p(a, b)$ be $\neq 0$.*[4]

We add two remarks: first of all, it is clear that the isomorphisms of $E_p(a, b)$ on W_1^* will not, in general, be algebraic ones; on the other hand, the assumption that K is algebraically closed cannot be removed; we have seen for instance in [4] that if K is the prime field of characteristic 5, the group $E_5(a, b)$ cannot be isomorphic to W_1^* if $4b^4 + a^3b^2 + 2a^6 \neq 0$.

10. Cores of abelian Lie algebras. In the two following sections, we keep the assumption that K is an *algebraically closed field.* Let G be an abelian formal Lie group of dimension n, and let $\mathfrak{g}_0$ be its Lie algebra. G being abelian, the structure of $\mathfrak{g}_0$ is entirely determined by the mapping $q: X \rightarrow X^p$ of the vector space $\mathfrak{g}_0$ into itself, which is *semilinear* with respect to the automorphism $\sigma: \xi \rightarrow \xi^p$ of K. The arguments which lead to Fitting's lemma [1, p. 132] apply as well here to the mapping q: let k be the smallest integer such that $q^k(\mathfrak{g}_0) = q^{k+1}(\mathfrak{g}_0)$; then k is also the smallest integer such that $q^{-k}(0) = q^{-(k+1)}(0)$; $\mathfrak{g}_0$ is the direct sum of the two Lie algebras $\mathfrak{h} = q^k(\mathfrak{g}_0)$ and $\mathfrak{k} = q^{-k}(0)$, the restriction of q to $\mathfrak{h}$ is a *one-one and onto mapping of* $\mathfrak{h}$, and for every $X \varepsilon \mathfrak{k}$, one has $X^{p^k} = 0$, k being the smallest integer having that property. We will say that $\mathfrak{h}$ is the *core* of the Lie algebra $\mathfrak{g}_0$.

---

[3] The differential $dx/v_1(x)$ is a differential of the first kind when x is taken as an uniformizing parameter; it is known that the fact that the coefficient of x^{p-1} in this differential vanishes, is independent of the choice of the uniformizing parameter (this follows also from Prop. 4 above), and that, by using Riemann-Roch's theorem, this condition is necessary and sufficient for the existence of a function in the field having a development of the type $x^{-p} + a_0 + a_1x + \cdots$. But this last condition is equivalent to the vanishing of the Hasse invariant [6] (I am indebted to A. Weil for these precisions).

[4] If we remark that, in the clasical case, an elliptic integral defines a (local) isomorphism of the elliptic group on the *additive* group, we may consider an isomorphism of $E_p(a, b)$ on W_1^* in the present case, as an analogon to an elliptic integral.

PROPOSITION 5. *If* $\mathfrak{h} \neq \{0\}$, *there exists a basis* $X_1, \cdots, X_m$ *of the core* $\mathfrak{h}$ *of* $\mathfrak{g}_0$ *such that* $X_i^p = X_i$ *for* $1 \leqq i \leqq m$.

We use induction on the dimension m of $\mathfrak{h}$. Let $Y \neq 0$ be an element of $\mathfrak{h}$, and let k be the smallest integer such that $Y, Y^p, \cdots, Y^{p^k}$ are linearly dependent; we have therefore

$$Y^{p^k} = \sum_{j=0}^{k-1} \lambda_j Y^{p^j} \tag{39}$$

and by definition of $\mathfrak{h}$, the λ_j cannot be all equal to 0. Let us determine an element $X_1 = \sum_{j=0}^{k-1} \xi_j Y^{p^j}$ by the condition $X_1^p = X_1$; as $Y, Y^p, \cdots, Y^{p^{k-1}}$ are linearly independent, this yields the equations

$$\begin{cases} \xi_0 = \lambda_0 \xi_{k-1}{}^p \\ \xi_1 = \xi_0{}^p + \lambda_1 \xi_{k-1}{}^p \\ \cdots\cdots\cdots\cdots \\ \xi_{k-1} = \xi_{k-2}{}^p + \lambda_{k-1} \xi_{k-1}{}^p \end{cases} \tag{40}$$

which determine completely $\xi_0, \xi_1, \cdots, \xi_{k-2}$ as functions of $t = \xi_{k-1}$, and give for t the equation

$$t = \lambda_0{}^{p^k} t^{p^{k+1}} + \lambda_1{}^{p^{k-1}} t^{p^k} + \cdots + \lambda_{k-1} t^p. \tag{41}$$

As the λ_j are not all 0, this equation has at least one root $\neq 0$ in K, which proves the existence of $X_1 \neq 0$ such that $X_1^p = X_1$. Let now $Y_2, \cdots, Y_m$ constitute with X_1 a basis of $\mathfrak{h}$; we can write

$$Y_i^p = \sum_{j=2}^{m} \beta_{ij} Y_j + \gamma_i X_1 \qquad (2 \leqq i \leqq m)$$

and the definition of $\mathfrak{h}$, together with the relation $X_1^p = X_1$, imply that the matrix (β_{ij}) of order $m-1$, is invertible. The inductive hypothesis proves therefore the existence of $m-1$ elements $T_2, \cdots, T_m$ which, together with X_1, constitute a basis for $\mathfrak{h}$, and are such that

$$T_i^p = T_i + \mu_i X_1 \qquad (2 \leqq i \leqq m).$$

These relations can also be written

$$(T_i + \xi_i X_1)^p = (T_i + \xi_i X_1) + (\xi_i^p - \xi_i + \mu_i) X_1 \qquad (2 \leqq i \leqq m).$$

It is therefore possible to choose the ξ_i $(2 \leqq i \leqq m)$ in K such that the elements $X_i = T_i + \xi_i X_1$ $(2 \leqq i \leqq m)$ constitute, together with X_1, a basis of $\mathfrak{h}$ satisfying the conditions of Proposition 5. The same argument may be applied to any Lie *subalgebra* of $\mathfrak{h}$; but if $Y = \sum_{i=1}^{m} \rho_i X_i$, the relation $Y^p = Y$

is equivalent to the m equations $\rho_i^p = \rho_i$, and means therefore that the coefficients ρ_i $(1 \leqq i \leqq m)$ belong to the *prime field* $\boldsymbol{F}_p$ of characteristic p. Hence:

PROPOSITION 6. *Let V be the vector space over $\boldsymbol{F}_p$ generated by $X_1, \cdots, X_m$. There is a one-to-one correspondence between the Lie subalgebras of $\mathfrak{h}$ and the subspaces of V; in particular, every subalgebra of $\mathfrak{h}$ is a direct summand.*

This shows in particular that the number of Lie subalgebras of $\mathfrak{h}$ is *finite*, and equal to the number of subspaces of V.

11. The structure of the Lie algebra $\mathfrak{k}$, supplementary to $\mathfrak{h}$ in $\mathfrak{g}_0$, is easily obtained by the Weyr-Fitting method [1]; let $s = n - m$ be the dimension of k, s_i $(1 \leqq i \leqq k)$ the dimension of $q^{i-1}(q^{-i}(0))$. The sequence $(s_i)_{1 \leqq i \leqq k}$ is decreasing, we have $s = \sum_{i=1}^{k} s_i$ and there is a basis (Y_{it}) of $\mathfrak{k}$ such that $1 \leqq i \leqq k$, $1 \leqq t \leqq s_i$ for each i, and which satisfies the following relations

$$(41) \qquad Y_{1t}{}^p = 0 \quad (1 \leqq t \leqq s_1), \qquad Y_{it}{}^p = Y_{i-1,t} \quad (2 \leqq i \leqq k, 1 \leqq t \leqq s_i).$$

We suppress the details of the proof. We will say that $\mathfrak{k}$ is the *p-radical* of the Lie algebra $\mathfrak{g}_0$.

12. Direct products. The notion of *direct product* of two formal Lie groups G_1, G_2 over K is defined in a natural way. Let n_1, n_2 be the dimensions of G_1 and G_2; it will be convenient to denote by $\alpha, \beta, \cdots$ the systems of n_1 indices, by $\lambda, \mu, \cdots$ the systems of n_2 indices, a system of $n_1 + n_2$ indices being written (α, λ); moreover, we will reserve the usual notations $x_\alpha, X_{hi}, Z_\alpha$, etc. to G_1, the corresponding notations for G_2 being $y_\lambda, Y_{hj}, T_\lambda$, etc. With these conventions, if the group laws of G_1 and G_2 are respectively

$$x_i'' = \phi_i(\boldsymbol{x}, \boldsymbol{x}') \qquad (1 \leqq i \leqq n_1)$$

$$y_j'' = \psi_j(\boldsymbol{y}, \boldsymbol{y}') \qquad (1 \leqq j \leqq n_2)$$

the group law of $G_1 \times G_2$ simply consists, by definition, in the juxtaposition of the preceding $n_1 + n_2$ formulas. It follows at once from this definition that, in the Taylor formula (2) of $G_1 \times G_2$, the operator of index α can be identified with Z_α (which means that it is the same polynomial in the X_{hi} as Z_α), the operator of index λ can be identified with T_λ; moreover, Z_α and T_λ commute, and the operator of index (α, λ) is their product $Z_\alpha T_\lambda$. We can express these facts by saying that the hyperalgebras of G_1 and G_2 are *typical subalgebras* [3, no. 18] of the hyperalgebra of $G_1 \times G_2$, and that the latter is the *tensor product* of these two hyperalgebras. Conversely, let G be a formal Lie group of dimension $n_1 + n_2$; let us write again α for a system of $n_1 + n_2$ indices, the last n_2 of which are 0, λ for a system of $n_1 + n_2$ indices, the first n_1 of which are 0, and let us adopt conventions similar as the previous ones. Then:

PROPOSITION 7. *Suppose the operators* X_α *form a typical subalgebra* $\mathfrak{L}$ *and the operators* Y_λ *a typical subalgebra* $\mathfrak{M}$ *of the hyperalgebra* $\mathfrak{G}$ *of* G, *and that* $X_\alpha Y_\lambda = Y_\lambda X_\alpha$ *for every pair* (α, λ). *Then, if the group law of* G *is canonical,* G *is isomorphic to the direct product* $G_1 \times G_2$ *of two groups having respectively* $\mathfrak{L}$ *and* $\mathfrak{M}$ *as hyperalgebras.*

Indeed, we know there exist two groups G_1, G_2 having canonical group laws and whose hyperalgebras are respectively $\mathfrak{L}$ and $\mathfrak{M}$ [5, Th. 1]. If we prove that the operator of index (α, λ) in the Taylor formula (2) of G is $Z_\alpha T_\lambda$, the proposition will result from the uniqueness theorem for groups having a given hyperalgebra [5, Th. 2]. Let $U_{(\alpha,\lambda)}$ be this operator, and let us use induction on $|\alpha| + |\lambda|$. Formula (6) then gives

$$U_{(\alpha,\lambda)}(fg) = f \cdot U_{(\alpha,\lambda)} g + g \cdot U_{(\alpha,\lambda)} f + \sum (T_\mu Z_\beta f)(T_{\lambda-\mu} Z_{\alpha-\beta} g)$$

where in the summation, we let the indices range over the set $0 \leqq \beta \leqq \alpha$, $0 \leqq \mu \leqq \lambda$, with the exception of the pairs of indices $(0, 0)$ and (α, λ). On the other hand

$$T_\lambda Z_\alpha(fg) = T_\lambda\left(\textstyle\sum_{0 \leqq \beta \leqq \alpha} (Z_\beta f)(Z_{\alpha-\beta} g)\right) = \sum_{\substack{0 \leqq \beta \leqq \alpha \\ 0 \leqq \mu \leqq \lambda}} (T_\mu Z_\beta f)(T_{\lambda-\mu} Z_{\alpha-\beta} g)$$

which shows that the difference $U_{(\alpha,\lambda)} - T_\lambda Z_\alpha$ is a derivation; but from the assumption that the group law is canonical, it follows that the preceding difference must be 0 if $|\alpha| + |\lambda| > 1$, and this ends our proof.

We can get rid of the assumption that the group law of G is canonical if we assume a little more about the subgroups of G:

PROPOSITION 8. *Suppose that the equations* $x_i = 0$ *for* $i \leqq n_1$ *define a subgroup* G_2 *of* G, *and that the equations* $x_i = 0$ *for* $n_1 < i \leqq n_1 + n_2$ *define another subgroup* G_1 *of* G. *Then, if* $X_\alpha Y_\lambda = Y_\lambda X_\alpha$ *for every pair* (α, λ), G *is isomorphic to the direct product* $G_1 \times G_2$.

Indeed, if we perform in the standard way the isomorphism $\boldsymbol{u}$ of G onto a group $\bar{G}$ with canonical law (no. 4), the images of G_1 and G_2 under $\boldsymbol{u}$ are defined by annihilating the coordinates of the same indices as those whose vanishing define G_1 and G_2 respectively ([3, no. 18] and, more correctly, [5, no. 3]); moreover, the $\bar{X}_\alpha$ (resp. $\bar{Y}_\lambda$) in the hyperalgebra of $\bar{G}$ are functions of the X_β alone (resp. of the Y_μ alone), hence $\bar{G}$ satisfies all the assumptions of Proposition 7, and the conclusion follows.

The preceding definitions and results extend immediately to the product of a finite number of formal Lie groups.

13. Cores of abelian Lie groups. We can now state the following structure theorem:

THEOREM 3. *Let G be an abelian formal Lie group over an algebraically closed field, $\mathfrak{g}_0$ its Lie algebra, $\mathfrak{h}$ the core and $\mathfrak{k}$ the p-radical of $\mathfrak{g}_0$. Then, if $\mathfrak{h}$ has dimension m, G is isomorphic to the direct product of m groups isomorphic to I_0 and of a group having $\mathfrak{k}$ as its Lie algebra.*

We first observe that a linear isomorphism $\boldsymbol{u}$ of a group G on a group $\bar{G}$, with $u_i(\boldsymbol{x}) = \sum_{j=1}^{n} \alpha_{ij}x_j$, has a derived isomorphism $\boldsymbol{u}'$ such that $\boldsymbol{u}'(X_{0i}) = \sum_{j=1}^{n} \alpha_{ji}\bar{X}_{0j}$. We can therefore always suppose that the m first variables (which we shall write x_i) are such that

$$X_{0i}^p = X_{0i} \qquad (1 \leqq i \leqq m) \tag{42}$$

whereas the $s = n - m$ remaining variables, written y_{jt} (notations of **11**) are such that the corresponding derivations Y_{0jt} in $\mathfrak{g}_0$ satisfy

$$Y_{01t}^p = 0 \quad (1 \leqq t \leqq s_1), \qquad Y_{0jt}^p = Y_{0,j-1,t} \tag{43}$$
$$(2 \leqq j \leqq k, 1 \leqq t \leqq s_j, \sum_{j=1}^{k} s_j = s).$$

In other words, the X_{0i} constitute a basis of $\mathfrak{h}$, the Y_{0jt} a basis of $\mathfrak{k}$. We adopt moreover the conventions of **12**, writing α for a system of n indices, the last $n - m = s$ of which are 0, and λ for a system of n indices, the first m of which are 0; an arbitrary system of n indices can therefore be denoted (α, λ); the operator of index α will be written Z_α, the operator of index λ will be written T_λ, and $U_{(\alpha,\lambda)}$ will denote the operator of index (α, λ). As a basis for an inductive argument, let us suppose that G satisfies in addition the following assumptions (A_r) $(r \geqq 0)$:

a) the group law of G is canonical;

b) $X_{hi}^p = X_{hi}$ for $0 \leqq h \leqq r$ and $1 \leqq i \leqq m$, and for every index $\alpha = (\alpha_1, \cdots, \alpha_m, 0, \cdots, 0)$ of height $< r$, $Z_\alpha = Z_{\alpha_1} Z_{\alpha_2} \cdots Z_{\alpha_m}$, where Z_{α_i}

is the same polynomial in $X_{0i}, X_{1i}, \cdots, X_{r-1,i}$ as the operator of the same index in the Taylor formula (2) of the group I_0;

c) for $0 \leqq h \leqq r$, $Y_{hjt}{}^p$ can be expressed as a polynomial in the Y_{ljt} $(l \leqq h)$ alone; and for every index λ of height $< r$, T_λ is a polynomial in the Y_{hjt} alone $(h < r)$.

We want to determine an isomorphism of G on a group which will satisfy assumptions (A_{r+1}); we proceed in several steps.

I. Let us prove first that for any index α *of height* r, Z_α is the same polynomial in the X_{hi} $(0 \leqq h \leqq r, 1 \leqq i \leqq m)$ as in the direct product of m groups identical to I_0. Using Prop. 7, we are reduced to proving that, for $p^{r-1}+1 \leqq l \leqq p^r - 1$, $Z_{l\epsilon_i}$ is equal to the polynomial $\phi_l(X_{0i}, \cdots, X_{r-1,i})$, where $\phi_l(X_0, \cdots, X_{r-1})$ is the polynomial equal to Z_l in the hyperalgebra of the group I_0. We notice that by assumption (A_r, b), $Z_{l\epsilon_i} = \phi_l(X_{0i}, \cdots, X_{r-1,i})$ for $0 \leqq l \leqq p^{r-1}$; let us argue by induction on $l > p^{r-1}$. It is clear that the "Leibniz formula" (6) for $Z'_l = \phi_l(X_{0i}, \cdots, X_{r-1,i})$ can be written

$$Z'_l(fg) = f \cdot Z'_l g + g \cdot Z'_l f + \sum (V_\beta f)(W_\gamma g) \tag{44}$$

where the operators V_β and W_γ are polynomials in $X_{0i}, \cdots, X_{r-1,i}$ which depend only on the "Leibniz formulae" for the X_{hi} $(0 \leqq h \leqq r-1)$ and on the expressions of $X_{hi}{}^p$ $(0 \leqq h \leqq r-1)$; as by assumption, these are the *same* as in I_0 (X_{hi} replacing X_h), the Leibniz formula (44) for Z'_l is deduced from that of Z_l (in the group I_0) by the same replacements; but then the inductive assumption shows that the difference $Z_{l\epsilon_i} - Z'_l$ is a derivation, and the assumption that the group laws of G and of I_0 are canonical proves finally that $Z_{l\epsilon_i} = Z'_l$.

II. We prove next that for every index λ *of height* r, T_λ is a polynomial in the Y_{hjt} alone $(h \leqq r)$. We use again the assumption that the result holds for the indices of height $< r$ (by (A_r, c)), and we use induction on the total degree $|\lambda|$. In the Leibniz formula for T_λ

$$T_\lambda(fg) = f \cdot T_\lambda g + g \cdot T_\lambda f + \textstyle\sum_{0 \leqq \mu \leqq \lambda} (T_\mu f)(T_{\lambda-\mu} g) \tag{45}$$

all the operators T_μ which intervene in the summation are polynomials in the Y_{hjt} alone. Suppose this was not true for T_λ; in the expression of T_λ there might then exist terms in the X_{hi} alone or "mixed" terms containing both the X_{hi} and Y_{hjt}. Let $T'_\lambda = \sum_\beta b_\beta X_\beta$ be the sum of the terms of the first type, $T''_\lambda = \sum_{\beta,\mu} b_{\beta\mu} X_\beta Y_\mu$ the sum of the terms of the second type, $T'''_\lambda = T_\lambda - T'_\lambda - T''_\lambda$ the sum of the terms in the Y_μ alone. Now, from the assumption that the group law is canonical, it follows that T'_λ cannot be a derivation if the b_β are not all 0, and therefore it follows from I that $T'_\lambda(fg) - f \cdot T'_\lambda g - g \cdot T'_\lambda f$ would then be a linear combination of terms $(X_\gamma f)(X_\delta g)$ with coefficients not all 0. Similarly $T''_\lambda(fg) - f \cdot T''_\lambda g - g \cdot T''_\lambda f$ would be a linear combination of terms $(Vf)(Wg)$ where V and W are monomials such that VW contains at least one X_{hi} *and* at least one Y_{hjt}, and the coefficients of these terms would not all be 0 if the $b_{\beta\mu}$ were not all 0. Finally, it is clear that the expression $T'''_\lambda(fg) - f \cdot T'''_\lambda g - g \cdot T'''_\lambda f$ is a sum of terms of type $(Y_\mu f)(Y_\nu g)$. If we compare these results with the right hand side in (45), we see that we must have $T'_\lambda = T''_\lambda = 0$, which proves our assertion (use being made, of course, of the linear independence of the operators X_α and Y_λ).

III. From I and the study of the group I_0 made in **7**, it follows that $Z_\alpha{}^p = Z_\alpha$ for all indices α of height $\leqq r$. Lemma 2 applied to $X_{r+1,i}$ $(1 \leqq i \leqq m)$ proves therefore that the difference $X_{r+1,i}{}^p - X_{r+1,i}$ is a *derivation*; in other words, we have

$$X_{r+1,i}{}^p = X_{r+1,i} + \sum_{i'} a_{ii'} X_{0i'} + \sum_{j',t'} b_{ij't'} Y_{0j't'}. \tag{46}$$

On the other hand, II and the assumption (A_r, c) show that for all indices λ of height $\leqq r$, $T_\lambda{}^p$ is a polynomial in the Y_{hjt} $(0 \leqq h \leqq r)$ alone; from this and Lemma 2 it follows that the difference

$$Y_{r+1,j,t}{}^p(fg) - f \cdot Y_{r+1,j,t}{}^p(g) - g \cdot Y_{r+1,j,t}{}^p(f)$$

is a sum of terms of type $(Y_\mu f)(Y_\nu g)$. The same argument as in II proves then that one has

$$Y_{r+1,j,t}{}^p = Y'_{jt} + \sum_{i'} e_{jti'} X_{0i'} \tag{47}$$

where Y'_{jt} is a polynomial in the $Y_{hj't'}$ $(h \leqq r+1)$.

We now define an isomorphism $\boldsymbol{u}$ of G on a group $\bar{G}$ such that the power series defining the inverse $\boldsymbol{u}^{-1}$ have the form

$$\begin{cases} u_i{}^{-1}(\bar{\boldsymbol{x}}) = \bar{x}_i + \sum_{i'} \xi_{ii'} \bar{x}_{i'}{}^{p^{r+1}} + \sum_{j',t'} \eta_{ij't'} \bar{y}_{j't'}{}^{p^{r+1}} \\ u_{jt}{}^{-1}(\bar{\boldsymbol{x}}) = \bar{y}_{jt} + \sum_{i'} \xi_{jti'} \bar{x}_{i'}{}^{p^{r+1}}, \end{cases} \tag{48}$$

the coefficients having to be determined later. This gives, for the derived homomorphism $(\boldsymbol{u}^{-1})' = (\boldsymbol{u}')^{-1}$, first

$$(\boldsymbol{u}^{-1})'(\bar{X}_{hi}) = X_{hi}, \qquad (\boldsymbol{u}^{-1})'(\bar{Y}_{hjt}) = Y_{hjt}$$

for $h \leqq r$, and further

$$(49) \qquad \begin{cases} (\boldsymbol{u}^{-1})'(\bar{X}_{r+1,i}) = X_{r+1,i} + \sum_{i'} \xi_{i'i} X_{0i'} + \sum_{j',t'} \xi_{j't'i} Y_{0j't'} \\ (\boldsymbol{u}^{-1})'(\bar{Y}_{r+1,j,t}) = Y_{r+1,j,t} + \sum_{i'} \eta_{i'jt} X_{0i'}. \end{cases}$$

If we take into account formulae (43), this gives first

$$(50) \qquad \begin{aligned} (\boldsymbol{u}^{-1})'(\bar{X}_{r+1,i}{}^{p} - \bar{X}_{r+1,i}) &= \sum_{i'} (\xi_{i'i}{}^{p} - \xi_{i'i} + a_{ii'}) X_{0i'} \\ &+ \sum_{j',t'} (b_{ij't'} - \xi_{j't'i}) Y_{0j't'} + \sum_{j',t'} \xi_{j't'i}{}^{p} Y_{0,j'-1,t'}. \end{aligned}$$

We can therefore choose first the $\xi_{ii'}$ such that in the right hand side the coefficients of the $X_{0i'}$ are all 0. We choose next the $\xi_{j't'i}$: for $j'=k$, $1 \leqq t' \leqq s_k$ we take $\xi_{kt'i} = b_{ikt'}$, and the coefficients of the $Y_{0kt'}$ in the right hand side of (50) vanish. Suppose next the $\xi_{j't'i}$ have been determined for all values $j' > l$; then, for $1 \leqq t' \leqq s_{l+1}$, the coefficient of $Y_{0lt'}$ in (50) is $b_{ilt'} - \xi_{lt'i} + \xi_{l+1,t',i}{}^{p}$; for $s_{l+1} < t' \leqq s_l$, the coefficient of $Y_{0lt'}$ is simply $b_{ilt'} - \xi_{lt'i}$; in both cases, it is possible to determine $\xi_{lt'i}$ such that these coefficients vanish. Relation (50) then gives

$$(51) \qquad \bar{X}_{r+1,i}{}^{p} = \bar{X}_{r+1,i} \qquad (1 \leqq i \leqq m).$$

To go further, we observe that in formula (47), Y'_{jt} is the sum of a linear combination of the $Y_{r+1,j',t'}$ and of a term Y''_{jt} which is of height $\leqq r$. Moreover, if we apply to both sides of (47) the iterated Frobenius homomorphism $\boldsymbol{p}'^{r+1}$, and take into account the relations (43), we see that we have

$$Y'_{jt} = Y_{r+1,j-1,t} + Y''_{jt}$$

(where, for $j=1$, $Y_{r+1,j-1,t}$ must be replaced by 0). We deduce from these remarks and from (49) that

$$(\boldsymbol{u}^{-1})'(\bar{Y}_{r+1,j,t}{}^{p} - \bar{Y}_{r+1,j-1,t}) = Y''_{jt} + \sum_{i'} (\eta_{i'jt}{}^{p} - \eta_{i',j-1,t} + e_{jti'}) X_{0i'}.$$

Here, we first determine the $\eta_{i'jt}$ for $j=1$, then by induction for the successive values of $j \leqq k$, in such a way that the coefficients of the $X_{0i'}$ on the right hand side of (52) all vanish. It follows finally that we have

$$(53) \qquad \bar{Y}_{r+1,j,t}{}^{p} = \bar{Y}_{r+1,j-1,t} + \bar{Y}''_{jt}$$

and the group law of $\bar{G}$ satisfies assumptions (A_{r+1}) with the possible exception of (A_{r+1}, a); but the same argument as in **7** shows that this condition also can be obtained by performing a new isomorphism which will not disturb conditions b) and c).

Having achieved the passage from (A_r) to (A_{r+1}), the proof of Theorem 3 is brought to a close by the same argument as at the end of **7**; this gives an isomorphism of G onto a group whose group law is canonical and whose hyperalgebra is the tensor product of m typical subalgebras identical to the hyperalgebra of I_0, and of another typical subalgebra, whose Lie algebra is identical to $\mathfrak{k}$. The final step consists in applying Prop. 7.

14. If a group G is the direct product of a certain number of groups identical with I_0 and of a group M having a "coreless" Lie algebra, we will say that the direct product L of the groups I_0 is the *core* of G, and that M is the *p-radical* of G. These definitions are justified by the fact that L and M are *intrinsically* characterized in G, in the following sense: if $\boldsymbol{u}$ is an automorphism of G, then $\boldsymbol{u}(L) = L$ and $\boldsymbol{u}(M) = M$, or, more precisely, the $u_i(\boldsymbol{x})$ contain only the variables corresponding to L (resp. M) when the index i is the index of such a variable. This follows from the more general result:

THEOREM 4. *Let $G_1 = L_1 \times M_1$, $G_2 = L_2 \times M_2$ be two abelian Lie groups, over an algebraically closed field, such that L_1, L_2 are products of groups identical with I_0, and M_1, M_2 have coreless Lie algebras. Then, if $\boldsymbol{u}$ is a homomorphism of G_1 into G_2, $\boldsymbol{u}$ maps L_1 into L_2 and M_1 into M_2.*

If $\mathfrak{L}_1$, $\mathfrak{L}_2$, $\mathfrak{M}_1$, $\mathfrak{M}_2$ are the hyperalgebras of L_1, L_2, M_1, M_2, it will be enough (from formula (13)) to prove that $\boldsymbol{u}'(\mathfrak{L}_1) \subset \mathfrak{L}_2$ and $\boldsymbol{u}'(\mathfrak{M}_1) \subset \mathfrak{M}_2$. Let $\mathfrak{h}_1$, $\mathfrak{h}_2$ be the cores, $\mathfrak{k}_1$, $\mathfrak{k}_2$ the p-radicals of the Lie algebras of G_1 and G_2. The fact that $\boldsymbol{u}'(Z^p) = (\boldsymbol{u}'(Z))^p$ and the definition of the core and of the p-radical (**10**) show first of all that $\boldsymbol{u}'(\mathfrak{h}_1) \subset \mathfrak{h}_2$, $\boldsymbol{u}'(\mathfrak{k}_1) \subset \mathfrak{k}_2$. Let us use conventions similar to those of the preceding numbers, writing $Z_\alpha^{(1)}$, $Z_\alpha^{(2)}$, $T_\lambda^{(1)}$, $T_\lambda^{(2)}$ for operators of $\mathfrak{L}_1$, $\mathfrak{L}_2$,, $\mathfrak{M}_1$, $\mathfrak{M}_2$ respectively. Suppose we have proved that $\boldsymbol{u}'(X_{hi}^{(1)})$ belongs to $\mathfrak{L}_2$ and $\boldsymbol{u}'(Y_{hjt}^{(1)})$ belongs to $\mathfrak{M}_2$ for all operators such that $h < r$; as $\boldsymbol{u}'$ is a homomorphism, it follows that $\boldsymbol{u}'(Z_\alpha^{(1)})$ belongs to $\mathfrak{L}_2$, and $\boldsymbol{u}'(T_\lambda^{(1)})$ belongs to $\mathfrak{M}_2$ for all indices α, λ of height $< r$. Now we observe that, as a consequence of formulae (53) and (43), for every element $T_\lambda^{(2)}$ of $\mathfrak{M}_2$, there is an integer e such that $(T_\lambda^{(2)})^{p^e} = 0$. On the other hand, we must have $(\boldsymbol{u}'(X_{ri}^{(1)}))^{p^e} = \boldsymbol{u}'(X_{ri}^{(1)})$ and this shows at once that $\boldsymbol{u}'(X_{ri}^{(1)})$ cannot contain terms in $T_\lambda^{(2)}$ nor "mixed" terms $Z_\alpha^{(2)} T_\lambda^{(2)}$. On the other hand, if we suppose that $\boldsymbol{u}'(Y_{rjt}^{(1)})$ is not in $\mathfrak{M}_2$, the inductive hypothesis, together with formula (14), shows, as in part II of the argument of **13**, that we would have

$$\boldsymbol{u}'(Y_{rjt}^{(1)}) = V_{jt}^{(2)} + \sum_i a_{ijt} X_{0i}^{(2)} \tag{54}$$

where $V_{jt}^{(2)}$ belongs to $\mathfrak{M}_2$. But then, raising both sides of (54) to a sufficiently high power p^e shows at once that $a_{ijt}=0$ for all indices, and our induction may therefore proceed, which proves the theorem.

15. We will not try to study the structure of "coreless" abelian Lie groups, where a great variety of cases seem possible. For groups without radical (over an algebraically closed field), on the contrary, not only is their structure determined by Theorem 3, but it is also possible to determine completely their homomorphisms. These groups, as follows from Theorem 3, are isomorphic to direct products $(W_1^*)^n$, and we have only therefore to determine the homomorphisms of a group $(W_1^*)^m$ into a group $(W_1^*)^n$. We introduce the following convention: for any rational *p-adic integer* $\zeta=\sum_{h=0}^{\infty}\nu_h p^h$ $(0\leqq\nu_h\leqq p-1)$ we write $(1+x)^\zeta$ for the power series

$$(1+x)^{\nu_0}(1+x^p)^{\nu_1}\cdots(1+x^{p^h})^{\nu_h}\cdots$$

which is obviously meaningful. We have then the following theorem:

THEOREM 5. *Every homomorphism* $\boldsymbol{u}$ *of a group* $(W_1^*)^m$ *into a group* $(W_1^*)^n$ *has the form*

$$1+u_i(\boldsymbol{x})=\prod_{j=1}^{m}(1+x_j)^{\zeta_{ij}} \qquad (1\leqq i\leqq n) \tag{55}$$

where the ζ_{ij} *are arbitrary p-adic integers* $(1\leqq i\leqq n, 1\leqq j\leqq m)$.

If $\boldsymbol{u}$ is such a homomorphism, it follows from the definition of a direct product that

$$1+u_i(\boldsymbol{x})=\prod_{j=1}^{m}(1+u_i(0,\cdots,0,x_j,0,\cdots,0))$$

and therefore we are immediately reduced to the case $m=n=1$ (this is merely the usual argument which determines the homomorphisms of modules which are direct sums of submodules, as "matrices" the elements of which are homomorphisms of the submodules into each other).

Let therefore u be an endomorphism of W_1^*, and suppose $u\neq 0$; let r be the smallest number h such that $u'(X_h)\neq 0$. From this definition it follows that $u'(Z_k)=0$ for $0<k<p^r$, and relation (14) shows therefore that $u'(X_r)$ is a *derivation*, in other words

$$u'(X_r)=\nu_r X_0, \qquad \nu_r\neq 0.$$

Raising both sides of this equation to the power p, and remembering that $X_h^p=X_h$ for every $h\geqq 0$, we see that ν_r is in the prime field $\boldsymbol{F}_p$, and therefore can be identified with an integer such that $0 < \nu_r \leqq p - 1$. If v_r is the automorphism of $W_1{}^*$ such that $1 + v_r(x) = (1 + x)^{\nu_r}$, then it is immediate that $(v_r{}^{-1}u)'(X_r) = X_0$. Let us now suppose that we have $u'(X_h) = X_{h-r}$ for $r \leqq h < s$; this means that u' coincides with the iterated Frobenius homomorphism $\boldsymbol{p}'^r$ on all X_h such that $h < s$, hence if $k < p^s, u'(Z_k) = 0$ if k is not a multiple of p^r, and if $k = p^r h$, $u'(Z_k) = Z_h$. This remark, together with relations (6) and (14), prove that $u'(X_s) - X_{s-r}$ is a derivation, in other words

$$u'(X_s) = X_{s-r} + \nu_s X_0.$$

Raising both sides to the power p shows that ν_s is in $\boldsymbol{F}_p$, and we identify it again with an integer such that $0 \leqq \nu_s \leqq p - 1$. Then, if v_s is the automorphism of $W_1{}^*$ such that $1 + v_s(x) = (1 + x)^{1+p^{s-r}\nu_s}$, it is immediate to see that $(v_s{}^{-1}u)'(X_h) = 0$ for $h < r$, $(v_s{}^{-1}u)'(X_h) = X_{h-r}$ for $r \leqq h < s + 1$.

The proof is then concluded by the usual inductive argument and "passage to the limit": the "infinite product" $v = v_r v_{r+1} \cdots v_s \cdots$ is meaningful and has the form $1 + v(x) = (1 + x)^{\zeta}$, where ζ is an invertible p-adic integer; moreover, the definition of the v_s gives $(v^{-1}u)'(X_h) = \boldsymbol{p}'^r(X_h)$ for every h; the endomorphisms $v^{-1}u$ and $\boldsymbol{p}^r$ therefore coincide, in other words $1 + u(x) = (1 + x)^{p^r \zeta}$, which ends the proof.

16. Some examples. The Lie algebra of the additive Witt group W_n is easily determined by induction on n, if we observe that the mapping $\boldsymbol{u}$ such that $u_1(\boldsymbol{x}) = 0$, $u_i(\boldsymbol{x}) = x_{i-1}$ for $2 \leqq i \leqq n$, is a homomorphism of W_n into itself, the image of W_n under $\boldsymbol{u}$ being isomorphic to W_{n-1} [7]. As $\boldsymbol{u}'(X_{0i}) = X_{0,i+1}$ for $1 \leqq i \leqq n - 1$, $\boldsymbol{u}'(X_{0n}) = 0$, it is easy to prove by induction that

$$X_{0i}{}^p = X_{0,i+1} \text{ for } 1 \leqq i \leqq n - 1,\ X_{0n}{}^p = 0. \tag{56}$$

Indeed, relations (56) for $i > 1$ follow from the isomorphism of $\boldsymbol{u}(W_n)$ with W_{n-1}; moreover, as $\boldsymbol{u}'(X_{01}{}^p - X_{02}) = 0$, one must have $X_{01}{}^p = X_{02} + \lambda X_{0n}$, and we are reduced to showing that $\lambda = 0$. But from the definition of W_n, it follows easily that

$$X_{01} = D_{01} + w_2 D_{02} + \cdots + w_n D_{0n}$$

where the w_i have no constant term, and $w_n = x_1{}^{p-1}h(x_1, \cdots, x_{n-1})$, h having no constant term if $n > 2$. As $X_{0n} = D_{0n}$, the coefficient of X_{0n} in $X_{01}{}^p$ can have no constant term if $n > 2$, hence must be 0. For $n = 2$, the relations (56) are proved by direct computation.

Relations (56) show therefore that W_n is a coreless abelian group. We intend to return to the study of that group in another paper. With regard to the multiplicative Witt group W_n^*, we have the following result: *W_n^* is isomorphic to the direct product $W_1^* \times W_{n-1}$.* To show this, we remark that the relations $x_2 = \cdot\cdot\cdot = x_n = 0$ define in W_n^* a subgroup obviously isomorphic to W_1^*. Now, the "inverse" in the group W_1^* is given by the series $\theta(x_1) = -x_1/(1+x_1)$; it follows that the mapping $\boldsymbol{u}$ defined by $u_1(\boldsymbol{x}) = 0$, and

$$u_i(\boldsymbol{x}) = \phi_i(\theta(x_1), 0, \cdot\cdot\cdot, 0, x_1, \cdot\cdot\cdot, x_n) \qquad (2 \leqq i \leqq n)$$

(where the ϕ_i define the group law of W_n^*) is a homomorphism of W_n^* on its subgroup V_{n-1} defined by $x_1 = 0$. From this (or directly from Proposition 8) it follows that W_n^* is isomorphic to the direct product $W_1^* \times V_{n-1}$. It remains to show that V_{n-1} is isomorphic to W_{n-1}; if we remember that the group law of V_{n-1} "formalizes" the multiplication law mod. $(1 + \mathfrak{p}^n)$ in the group $1 + \mathfrak{p}$ of an unramified $\mathfrak{p}$-adic field [7, pp. 133-134], an explicit isomorphism of V_{n-1} on W_{n-1} will be given by the *p-adic logarithm,* where all operations in the power series are to be understood in the "vector" sense of Witt (including division by the denominators k, which are not ordinary integers, but sums—in the vector sense—of k unit vectors). We do not go into details.

Another interesting type of formal Lie groups stems from the Poincaré-E. Cartan method of defining "bilinear groups" associated with any (associative) algebra A with unit over a field K. Suppose $(a_i)_{1 \leq i \leq n}$ is a basis of A over K, with a_1 the unit element, and let $a_i a_j = \sum_k \gamma_{ijk} a_k$ be the multiplication table ($\gamma_{ijk} \varepsilon K$), with $\gamma_{1jk} = \delta_{jk}$, $\gamma_{i1k} = \delta_{ik}$. If we write the multiplication formula

$$(1 + z_1)a_1 + \textstyle\sum_{i>1} z_i a_i = ((1 + x_1)a_1 + \sum_{i>1} x_i a_i)((1 + y_1)a_1 + \sum_{i>1} y_i a_i)$$

we get a formal Lie group $G(A)$ over K:

$$(57) \qquad \begin{cases} z_1 = x_1 + y_1 + x_1 y_1 + \sum_{j>1, k>1} \gamma_{jk1} x_j y_k \\ z_i = x_i + y_i + \sum_{j>1} (x_1 y_j + y_1 x_j) + \sum_{j>1, k>1} \gamma_{jki} x_j y_k \qquad (i > 1) \end{cases}$$

If we extend the field of scalars to an overfield K', the group $G(A_{K'})$ corresponding to the extended algebra (and to the same basis (a_i)) is defined by the same formulae (57), but here isomorphisms defined by power series with coefficients in K' are of course permitted.

Let us consider only a few simple cases. Suppose first that A is a *separable* overfield of K. Then, if Ω is the algebraic closure of K, it is well known that A_Ω splits in the direct sum of n fields isomorphic with Ω, and therefore $G(A_\Omega)$ will be isomorphic to $(W_1{}^*)^n$, from which it follows at once that the abelian group $G(A)$ has *no p-radical.* On the other hand, suppose A is an inseparable extension generated by a single element θ, satisfying the irreducible equation $\theta^{p^e}=\alpha$ over K. Then in the field Ω, we have $\alpha=\beta^{p^e}$, and therefore the algebra A_Ω has here a basis over Ω consisting of $1, a=\theta-\beta, a^2, \cdots, a^{p^e-1}$, and we have $a^{p^e}=0$. It is immediate that the corresponding group $G(A_\Omega)$ is isomorphic to the group denoted by J in [2, no. 5], with $l=p^e$; hence $G(A)$ is here a *coreless* abelian group. These two examples justify, in some way, the term "p-radical" introduced above and the method outlined here may perhaps be of some use in the study of associative algebras over arbitrary fields.

NORTHWESTERN UNIVERSITY.

BIBLIOGRAPHY.

[1] J. Dieudonné, "Sur la réduction canonique des couples de matrices," *Bull. Soc. Math. France*, vol. 74 (1946), pp. 130-146.

[2] ———, "Sur les groupes de Lie algébriques sur un corps de caractéristique $p>0$," *Rend. Circ. Mat. Palermo*, (2), vol. 1 (1952), pp. 380-402.

[3] ———, "Groupes de Lie et hyperalgèbres de Lie sur un corps de caractéristique $p>0$," *Comm. Math. Helv.*, vol. 28 (1954), pp. 87-118.

[4] ———, "Sur quelques groupes de Lie abéliens sur un corps de caractéristique $p>0$," *Archiv der Math.*, vol. 5 (1954), pp. 274-281.

[5] ———, "Sur la notion de variables canoniques," *Anais Acad. Bras. de Ciencias*, (1955).

[6] H. Hasse, "Existenz separabler zyklischer unverzweigter Erweiterungskörper vom Primzahlgrade p über elliptischen Funktionenkörpern der Charakteristik p," *J. reine und angew. Math.*, vol. 172 (1934), pp. 77-85.

[7] E. Witt, "Zyklische Körper und Algebren der Charakteristik p vom Grad p^n," *J. reine und angew. Math.*, vol. 176 (1937), pp. 126-140.

DOUBLY STOCHASTIC MATRICES AND COMPLEX VECTOR SPACES.*

By SEYMOUR SHERMAN.

A doubly stochastic (d. s.) matrix is a matrix P such that $P_{ij} \geqq 0$, $\sum_i P_{ij} = \sum_j P_{ij} = 1$ for all i and j. A. Horn has proved

THEOREM 1. *If $y = Px$, where x, y are complex n-vectors, and P is a* d. s. *matrix, and $c_1, c_2, \cdots, c_n$ are any complex numbers, then $\sum_{i=1}^{n} c_i y_i$ lies in the convex hull of all the points $\sum_{i=1}^{n} c_i x_{\alpha i}$, $\alpha \varepsilon R^n$, where R^n is the set of all the permutations of $(1, \cdots, n)$*

and conjectured the truth of

THEOREM 2. *If x, y are complex n-vectors and $c_1, c_2, \cdots, c_n$ are any complex numbers imply that $\sum_{i=1}^{n} c_i y_i$ lies in the convex hull of the vectors $\sum_{i=1}^{n} c_i x_{\alpha i}$, $\alpha \varepsilon R^n$, then $y = Px$ where P is a* d. s. *matrix.*

In what follows Theorem 2 is established.

Let E be complex n-space. Let η represent the general complex linear functional on E $(\eta \varepsilon E^*)$ and the value of η for some $x \varepsilon E$ is represented by (η, x). If we consider E as real $2n$-space, then each real linear functional ρ on E has the property that for some $\eta \varepsilon E^*$ $(\rho, x) = R(\eta, x)$, where $R(\eta, x)$ is the real part of (η, x).

LEMMA 1. *Let X be a compact convex set in E. Suppose that for each $\eta \varepsilon E^*$ $(\eta, y) \varepsilon (\eta, X) = \{(\eta, z) : z \varepsilon X\}$. Then $y \varepsilon X$.*

Proof. Since $(\eta, y) \varepsilon (\eta, X)$, it follows $R(\eta, y) \varepsilon R(\eta, X)$. But then from a standard separation theorem ([2], p. 47) it follows that $y \varepsilon X$.

If Lemma 1 is applied to the case where X is the convex hull of the vectors which are derived from x by taking all permutations of the components of x relative to a fixed complex coordinate system, then $y \varepsilon X$ for y satisfying the hypothesis of Theorem 2. Now note

---

* Received November 5, 1954.

LEMMA 2. *Let G be a finite collection $\{G_1, G_2, \cdots, G_n\}$ of linear transformations $E \to E$. Let $x \varepsilon E$. Denote by $K(G)$ the convex hull of G. Denote by $K(x)$ the convex hull of $Gx = \{G_1x, G_2x, \cdots, G_nx\}$. If $y \varepsilon K(x)$, then $y = Dx$ where $D \varepsilon H(G)$.*

Proof. Since $y \varepsilon K(x)$ it follows that $y = \sum_{i=1}^{n} w_i(G_ix)$, with $w_i \geqq 0$, $\sum w_i = 1$, and so $y = Dx$ with $D = \sum w_iG_i \varepsilon H(G)$. (There are extensions to the case where G is not finite but $K(x)$ is compact; since such results are not needed in the sequel they are not presented here.) The application of Lemma 2 to $y \varepsilon X$ implies that $y = Dx$ with x, y elements of the complex vector space E and D an n by n d. s. matrix (since D is a convex combination of permutation matrices). This establishes Theorem 2.

UNIVERSITY OF PENNSYLVANIA.

REFERENCES.

[1] A. Horn, "Doubly stochastic matrices and the diagonal of a rotation matrix," *American Journal of Mathematics*, vol. 76 (1954), pp. 620-630.

[2] W. Fenchel, "Convex cones, sets and functions," Princeton University Logistics Research Project, September, 1953.

ON COVARIANT FIBERINGS OF KLEIN SPACES.*[1]

By G. D. Mostow.

1. Introduction.

By a Klein space we mean a space on which a Lie group of transformations operates transitively; or what is equivalent, the factor space G/S of a Lie group G by a closed subgroup S. If the group G is compact then, of course, G/S is compact, but not conversely. In case G is compact, we call the Klein space G/S compact-by-heredity or "*h*-compact." This paper is devoted to investigating the relation between general Klein spaces and *h*-compact spaces.

In the special case that the isotropy group (i. e., the set of transformations keeping a point fixed) consists of the identity alone, the Klein space is a group manifold and it is known that a connected Lie group is topologically a direct product of a compact subgroup and a Euclidean space. This fact raises the question: to what extent does a similar theorem hold for general Klein spaces? Examples show readily that such a result does not hold for Klein spaces when the isotropy subgroup is not connected. When, however, the isotopy subgroup is connected, there is a similar relation. For we can prove (cf. Theorem 3.1) the theorem that a Klein space H with connected isotropy group can be retracted by a strong deformation retraction to a subspace that is homeomorphic to an *h*-compact Klein space. Nevertheless, it is false that a general Klein space with connected isotropy group is a direct product of a compact subspace and a Euclidean space. This is shown by an example due to Samelson (cf. Section 5).

Instead, the kind of decomposition that seems to occur is that of *covariant fibering.* If H is a Klein space with associated group G, we mean by a G-covariant fibering a decomposition of the differentiable manifold H into Euclidean fibers such that 1) the fibers are permuted transitively by some maximal compact subgroup M of G and 2) the subgroup of M which keeps some fiber F invariant is equivalent to a linear group on F. Under these circumstances, some orbit of M is a cross-section set to the fibering

---

* Received February 19, 1954; revised October 18, 1954.

[1] Research performed in part under contract DA-36-034-ORD-1274 with Office of Ordnance Research, U. S. Army.

and it is of course an h-compact subspace to which H can be retracted via a strong deformation retraction along the fibers.

We devote this paper to the existence and uniqueness of covariant fiberings of Klein spaces. Our main results are:

THEOREM. *Any two G-covariant fiberings of a Klein space are equivalent, in the sense that there exists a one-to-one bundle map of one onto the other* (Section 7).

THEOREM 4.1. *A Klein space with associated group G whose isotropy subgroup is connected and self-adjoint modulo the radical of G admits a covariant fibering.*

Theorem 4.1 applies when the connected isotropy subgroup is either semi-simple, or compact, or in the radical. In particular, the theorem holds for solvmanifolds.

Theorem 4.1 depends on a new decomposition theorem for matrices that was derived by the author in [10]. In all probability the self-adjoint hypothesis can be dropped. This would entail proving a generalization of the cited decomposition theorem for matrices.

Our method for proving the existence of covariant fiberings consists in deriving decompositions for Lie groups, some of which have independent interest. For example:

THEOREM 2.2. *Let G be a connected Lie group. Then it contains a maximal compact subgroup M and a Euclidean subspace U such that $G = M \cdot U$ topologically, and $mUm^{-1} \subset U$ for all m in M.*

As a result of this theorem, we obtain the following sharpened form of the theorem on conjugacy of compact subgroups.

THEOREM 2.3. *Let M_1 and N_1 be compact subgroups of a connected Lie group G and let θ_1 be an inner automorphism of G which sends M_1 onto N_1. Let M and N be arbitrary maximal compact subgroups of G which include M_1 and N_1 respectively. Then there is an inner automorphism θ sending M onto N and coinciding with θ_1 on M_1.*

As an application of our results on covariant fiberings we prove the

THEOREM. *An aspherical Klein space is homeomorphic to Euclidean space.*

This result can be generalized to obtain a characterization of Euclidean space as an aspherical space whose group of auto-homeomorphisms contains a locally compact C_σ transitive subgroup (to appear).

2. Some decompositions for Lie groups.

Notation. Lie groups and their Lie algebras are denoted by corresponding Roman and German letters. $\mathfrak{G}$, the Lie algebra of G, is identified with the tangent space to G at the identity.

Let $\mathfrak{L}$ be a real semi-simple Lie algebra, i. e., one on which the Killing form $\operatorname{Tr} \operatorname{ad} X \operatorname{ad} Y$ is non-degenerate; or equivalently, one having no abelian ideals. By the adjoint group of $\mathfrak{L}$ we mean the connected Lie group of automorphisms of $\mathfrak{L}$ whose Lie algebra is $\operatorname{ad} \mathfrak{L}$. A subalgebra $\mathfrak{K}$ is called a compact subalgebra of $\mathfrak{L}$ if the analytic subgroup of the adjoint group with Lie algebra $\operatorname{ad} \mathfrak{K}$ is compact. It is known that if $\mathfrak{K}$ is a maximal compact subalgebra of $\mathfrak{L}$, then it has a unique orthogonal complement $\mathfrak{E}$ with respect to the Killing form (cf. [9]); moreover, if ρ is any representation of $\mathfrak{L}$ by linear transformations of a real linear space V, there is a base for V with respect to which $\rho(\mathfrak{K})$ and $\rho(\mathfrak{E})$ consist of skew-symmetric and symmetric matrices respectively (cf. [10]). To each maximal compact subalgebra $\mathfrak{K}$ there corresponds a unique automorphism $\theta_{\mathfrak{K}}$ of order two such that $\theta_{\mathfrak{K}}(X) = X$ or $-X$ according as X is in $\mathfrak{K}$ or $\mathfrak{E}$. We call $\theta_{\mathfrak{K}}$ the "$\mathfrak{K}$-star" automorphism of $\mathfrak{L}$.

Definition. A subalgebra of a real semi-simple algebra $\mathfrak{L}$ is called *self-adjoint in* $\mathfrak{L}$ if it is invariant under some star automorphism of $\mathfrak{L}$.

To say that the subalgebra $\mathfrak{C}$ is invariant under the $\mathfrak{K}$-star automorphism of $\mathfrak{L}$ is equivalent to $\mathfrak{C} = \mathfrak{C} \cap \mathfrak{K} + \mathfrak{C} \cap \mathfrak{E}$ where $\mathfrak{E}$ is the orthogonal complement of $\mathfrak{K}$. The author has proved in [10] that a semi-simple subalgebra of a semi-simple $\mathfrak{L}$ is self-adjoint in $\mathfrak{L}$.

Let L be a connected semi-simple group, C an analytic subgroup. We say that C is self-adjoint in L if the Lie subalgebra $\mathfrak{C}$ is self-adjoint in $\mathfrak{L}$. If G is a connected Lie group, C is an analytic subgroup, and R is the radical of G, we say that C is *self-adjoint modulo the radical* if CR/R is self-adjoint in the semi-simple group G/R.

It is convenient to know that a star automorphism of the Lie algebra of a semi-simple analytic group can be extended to an automorphism of the group. For let D be the center of the simply connected group L^* with semi-simple Lie algebra $\mathfrak{L}$, and let $\mathfrak{K} + \mathfrak{E}$ be a Cartan decomposition of $\mathfrak{L}$. Then D is contained in the analytic subgroup determined by K (cf. [9]). Now L^* being simply connected, the $\mathfrak{K}$-star automorphism of $\mathfrak{L}$ extends to an automorphism θ of L^*, and by the foregoing θ keeps the points of D fixed.

Hence θ induces an automorphism of any analytic group L locally isomorphic with L^*—we call it the K-star homomorphism of L, where K is the analytic subgroup of L determined by $\mathfrak{K}$.

It is clear that C is self-adjoint in L if and only if it is invariant under some star automorphism of the group L.

LEMMA 2.1. *Let $\mathfrak{C}$ be a self-adjoint subalgebra of the semi-simple Lie algebra $\mathfrak{L}$ and let $\mathfrak{R}$ be the radical of $\mathfrak{C}$. Let θ be a star-automorphism of $\mathfrak{L}$ which keeps $\mathfrak{C}$ invariant and let θ' denote the induced automorphism of $\mathfrak{C}/R$. Then θ' is a star-automorphism of $\mathfrak{C}/R$.*

Proof. Let θ be the $\mathfrak{K}$-star automorphism. Then $\mathfrak{L} = \mathfrak{K} + \mathfrak{E}$ and furthermore, coordinates may be selected in $\mathfrak{L}$ so that the matrix of ad X is skew-symmetric if $X \varepsilon \mathfrak{K}$, symmetric if $X \varepsilon \mathfrak{E}$ (cf. [9] or [10]). Now $\theta(\mathfrak{C}) = \mathfrak{C}$ implies $\mathfrak{C} = \mathfrak{C} \cap \mathfrak{K} + \mathfrak{C} \cap \mathfrak{E}$ and hence ad $\mathfrak{C}$ (considered as a subset of ad $\mathfrak{L}$) is represented by a self-adjoint family of matrices. As a result ad $\mathfrak{C}$ is a completely reducible family. Operating with ad $\mathfrak{C}$ on the radical $\mathfrak{R}$, we find that $\mathfrak{R}$ is a direct sum of $\mathfrak{C}$-ideals and is thus abelian. Operating with ad $\mathfrak{C}$ on $\mathfrak{C}$, we find that $\mathfrak{R}$ admits a complementary ideal and thus $\mathfrak{R}$ is central in $\mathfrak{C}$. It follows at once that $\mathfrak{C} = [\mathfrak{C}, \mathfrak{C}] + \mathfrak{R}$ (direct), and $[\mathfrak{C}, \mathfrak{C}]$ may be identified with the semi-simple quotient $\mathfrak{C}/\mathfrak{R}$. Let us denote $[\mathfrak{C}, \mathfrak{C}]$ by $\mathfrak{M}$. $\mathfrak{M}$ is invariant under θ and the automorphism θ' can obviously be identified with the restriction of θ to $\mathfrak{M}$. Now $\mathfrak{M} = \mathfrak{M} \cap \mathfrak{K} + \mathfrak{M} \cap \mathfrak{E}$ and $\theta'(X)$ is X if $X \varepsilon \mathfrak{M} \cap \mathfrak{K}$ and $-X$ if $X \varepsilon \mathfrak{M} \cap \mathfrak{E}$. Furthermore, $\mathfrak{M} \cap \mathfrak{K} + \mathfrak{M} \cap \mathfrak{E}$ is a Cartan decomposition of $\mathfrak{M}$ (cf. [10]). Consequently θ' is a star automorphism.

Let Γ be a group of automorphisms of a Lie group G and let S be a subset of G. Let $d\Gamma$ denote the set of differentials at the identity of the transformations of Γ, that is, the automorphisms of $\mathfrak{G}$ induced by Γ.

Definition. S is a Γ-*invariant exp-set* if there exist linearly independent subspaces $\mathfrak{S}_1, \cdots, \mathfrak{S}_n$ of $\mathfrak{G}$ which are invariant under $d\Gamma$ such that the mapping $s_1 + s_2 + \cdots + s_n \rightarrow \exp s_1 \cdot \exp s_2, \cdots, s_n$ is a homeomorphism of $\mathfrak{S}_1 + \cdots + \mathfrak{S}_n$ onto S (s_i in S_i). Γ is called completely reducible if $d\Gamma$ is a completely reducible (linear) group.

It should be noticed that a Γ-invariant exp-set S is invariant under Γ and homeomorphic to Euclidean space; also that the operation of Γ on S is equivalent to the operation of $d\Gamma$ on a linear subspace of $\mathfrak{G}$. These facts result from the well-known identity $T(\exp X) = \exp dT(X)$, for X in $\mathfrak{G}$. These facts result from the well-known identity $T(\exp X) = \exp dT(X)$, for

X in $\mathfrak{G}$ and any automorphism T, where dT denotes the differential of T at the identity.

If M is a subgroup of a Lie group G and Γ_M denotes the set of inner automorphisms of G by elements of M, we shall mean by an *M-invariant exp-set* a Γ_M-invariant exp-set.

LEMMA 2.2. *Let G be a connected Lie group containing no non-trivial central toroidal subgroup. Let $\mathfrak{N}$ be the maximum nilpotent ideal of the radical $\mathfrak{R}$. Then N is a closed simply connected subgroup of G.*

Proof. It is clear that the radical R is a closed subgroup of G. Inasmuch as N is the connected component of the identity in the set of all elements x of R such that $\operatorname{Ad} x$ has only 1 as eigenvalue ($\operatorname{Ad} x$ being the differential at the identity I of the inner automorphism $g \rightarrow xgx^{-1}$), it is clear that N is closed. For any non-zero X in $\mathfrak{N}$, $\exp X \neq I$. Otherwise $\operatorname{ad} X$ which is nilpotent on G and satisfies $\exp X = \operatorname{Ad} \exp X = 1$ must be equal to $\log \exp \operatorname{ad} X = \log 1 = 0$; that is, X is central in G and the one parameter subgroup $\{\exp tX \mid 0 \leqq t \leqq 1\}$ is a compact toroid, contrary to our assumption.

Now let N denote the simply connected covering space of N, let D be the kernel of the natural map of N onto N, and let us identify the Lie algebra of N with N. Since any element of a nilpotent group lies on a one-parameter subgroup (this is well-known, cf. [7] or [11]) for any x in D, there is an X in N such that $\exp X = x$ is in D. Since $\exp X = I$ in N, we conclude $X = 0$, $x = I$, $D = (I)$, and N is simply connected.

Definition. Let G be a topological group and let $A_1, \cdots, A_n$ be subspaces. We say that $G = A_1 \cdot A_2 \cdots A_n$ topologically if the mapping $\theta\colon (a_1, \cdots, a_n) \rightarrow a_1 a_2 \cdots a_n$ of $A_1 \times \cdots \times A_n$ into G is a homeomorphism onto, each A_i being given its relative topology with respect to G.

PROPOSITION 2.1. *Let G be an analytic group with non-trivial radical R and let C be a closed analytic subgroup. Assume that for every non-trivial abelian normal analytic subgroup V the topological closure of CV is G. Then there is a normal abelian analytic subgroup W such that $G = CW$.*

Proof. For typopgraphical convenience, we denote the topological closure of a set S by $\mathcal{C}\ell\, S$.

We let G^* denote the simply connected covering group of G, f the covering homomorphism and D^* the kernel of f. For any analytic subgroup S of G we denote by S^* the connected component of the identity of $f^{-1}(S)$ and vice versa, and we denote the Lie subalgebra of both by $\mathfrak{S}$. We denote by $[S, S]$ the commutator subgroup of S, it being the analytic subgroup with Lie algebra $[\mathfrak{S}, \mathfrak{S}]$.

From the hypotheses of the proposition it follows directly that for any non-trivial normal analytic abelian subgroup V (such a V exists since G is not semi-simple), $G = \mathcal{C}\ell\, D^*C^*V^*$. Since

$$[\mathcal{C}\ell\, D^*C^*V^*, \mathcal{C}\ell\, D^*C^*V^*] \subset \mathcal{C}\ell\, [D^*C^*V^*, D^*C^*V^*] \subset \mathcal{C}\ell\, [C^*V^*, C^*V^*],$$

we conclude that $[G^*, G^*] \subset \mathcal{C}\ell\, [C^*V^*, C^*V^*]$.

Now C^*V^* is a normal analytic subgroup of G^* since its Lie algebra is clearly invariant under $\mathrm{Ad}\, G^*$ and hence an ideal; hence $[C^*V, C^*V^*]$ is a normal analytic subgroup of the simply connected Lie group G^*. Hence $[C^*V^*, C^*V^*]$ is closed (cf. Pontriagin, *Topological Groups,* p. 279). Thus $[G^*, G^*] \subset [C^*V^*, C^*V^*]$ (cf. [7]). It follows now from the fact that a semi-simple Lie algera is its own commutator subalgebra that $G^*/R^* = [G^*, G^*]R^*/R^* \subset C^*R^*/R^*$. Hence $G^* = C^*R^*$. Letting S^* denote the radical of C^* and $L^* \cdot S^*$ a Levi decomposition of C^*, we conclude that $G^* = L^* \cdot R^*$ is a Levi decomposition of G^*. Inasmuch as G^* is simply connected and admits as a covering group a semi-direct product of L^* and R^*, we conclude that $G^* = L^* \cdot R^*$ (semi-direct) and in particular $L^* \cdot R^*$ is a direct product fibering of G^*.

We now set $D_L^* = D^*R^* \cap L^*$, $D_R^* = D^*L^* \cap R^*$. Since $x \rightarrow xR^* \cap L^*$ is a homomorphism of G^* onto L^*, D_L^* is central in L^*; since $d \rightarrow dL^* \cap R^*$ is a homomorphism of D^* into R^*, D_R^* is an abelian subgroup of R^*. Furthermore, from the well-known fact that D^* is a finitely generated group, it follows that D_R^* is finitely generated though it is, of course, not necessarily closed. Let $\rho(x)$ denote the restriction to the Lie algebra of R^* of $\mathrm{Ad}\, x$ for x in G^*. Then $\rho(L^*)$ is a semi-simple Lie group of matrices and it is known (cf. Mostow, *Annals of Mathematics,* vol. 52 (1950), p. 615) that its center is finite. Now for any r in D_R^*, there is an l in D_L^* such that $lv \,\varepsilon\, D^*$, and an n such that $\rho(l^n) = I$, the identity. Since $lrl^{-1} = l^{-1}(lr) = r$, l and r commute. Consequently, $\rho(r^n) = \rho(l^n)\rho(r^n) = \rho(l^n r^n) = \rho((lr)^n) = I$ and r^n is central in R. Since D_R^* is finitely generated, it contains a subgroup D_R' of finite index which is central in R. In particular, $\mathcal{C}\ell\, D_R'M$ and $\mathcal{C}\ell\, D_RM$ have the same connected component of the identity for any analytic subgroup M of R^*.

Now we have for any normal analytic abelian subgroup V (which is of necessity in the radical R),

$$G^* = \mathcal{C}\ell\, D^*C^*V^* = \mathcal{C}\ell\, L^* \cdot D_R^*S^*V^* = L^* \cdot \mathcal{C}\ell\, D_R^*S^*V^*,$$

since $L^* \cdot R^*$ is a direct product fibering. Hence $R^* = \mathcal{C}\ell D_R{}^* S^* V^*$ and R^* being connected, $R^* = \mathcal{C}\ell D_R{}' S^* V^*$.

Case 1. The center Z^* of R^* is not discrete.

Here Z^* contains a non-trivial abelian normal subgroup of G^* and hence $R^* = \mathcal{C}\ell D_R{}' S^* Z^* = \mathcal{C}\ell S^* Z^*$. Hence S^* is normal in R^* and being invariant under conjugations by elements of L^*, we conclude that S^* is normal in G^*. If $S^* =$ (identity), then $R^* = Z^*$, R is an abelian normal analytic subgroup and $G = CR$. If S^* is not trivial, it contains a characteristic abelian normal non-trivial subgroup V^*, namely, the last non-trivial subgroup in the sequence of successive commutator subgroups. Then $V = f(V^*)$ is normal in G and $G = \mathcal{C}\ell CV = \mathcal{C}\ell C = C$. Thus, the proposition is true in Case 1 ($W = R$ or $W =$ (identity)).

Case 2. The center of R^* is discrete.

Here $\mathfrak{R}$, the Lie algebra of R^*, has zero as center and hence it is not nilpotent. We denote by $\mathfrak{R}^2$ the derived algebra $[\mathfrak{R}, \mathfrak{R}]$ and recursively set $\mathfrak{R}^n = [\mathfrak{R}^{(n-1)}, \mathfrak{R}]$ and $\mathfrak{R}^\infty = \bigcap_n \mathfrak{R}^n$. $\mathfrak{R}^\infty = \mathfrak{R}^N$ for some N, and $\mathfrak{R}$ being non-nilpotent, $\mathfrak{R}^\infty \neq 0$. We denote by $\mathfrak{R}^{(n)}$ the ideal $[\mathfrak{R}^{(n-1)}, \mathfrak{R}^{(n-1)}]$ where $\mathfrak{R}^{(2)} = \mathfrak{R}^2$.

Case 2a. $[\mathfrak{R}, \mathfrak{R}]$ is not abelian.

Here $\mathfrak{R}^{(3)}$ is a non-zero characteristic ideal of $\mathfrak{G}$, the Lie algebra of G. Hence it contains a non-zero abelian ideal $\mathfrak{V}$ of G, whose corresponding subgroup in G^* we denote by V^*. From $R^* = \mathcal{C}\ell D_R{}' S^* V^*$ we infer

$$[R^*, R^*] = \mathcal{C}\ell\,[S^* V^*, S^* V^*] = [S^* V^*, S^* V^*],$$

the last assertion following since a normal analytic subgroup of a simple connected group is closed. Letting S denote the Lie algebra of S^*, we get

$$\text{(A)} \qquad [\mathfrak{R}, \mathfrak{R}] = [\mathfrak{S}, \mathfrak{S}] + [\mathfrak{S}, \mathfrak{V}] = [\mathfrak{S}, \mathfrak{S}] + [\mathfrak{R}, \mathfrak{R}]^2.$$

Now it is a well-known consequence of Lie's theorem on solvable Lie algebras that $[\mathfrak{R}, \mathfrak{R}]$ is a nilpotent Lie algebra. From equation (A) it follows readily that the subalgebra $[\mathfrak{S}, \mathfrak{S}]$ must coincide with the nilpotent $[\mathfrak{R}, \mathfrak{R}]$. Hence $[\mathfrak{R}, \mathfrak{R}] \subset \mathfrak{S}$, and C contains a non-trivial abelian normal analytic subgroup of G, namely, the smallest non-trivial $[R, R]^{(n)}$. Hence $G = \mathcal{C}\ell CV = \mathcal{C}\ell C = C$, and the proposition holds in Case 2a.

Case 2b. $[\mathfrak{R}, \mathfrak{R}]$ is abelian.

Let $\mathfrak{C}$ denote the Lie subalgebra of C^*. Since $R^* = \mathcal{C}\ell D_R{}' S^* [R^*, R^*]$,

$S^* \cap [R^*, R^*]$ is clearly a normal subgroup of R^*. Thus $\mathfrak{S} \cap [\mathfrak{R}, \mathfrak{R}] = \mathfrak{C} \cap [\mathfrak{R}, \mathfrak{R}]$ is an abelian ideal of $\mathfrak{G}$ in this case. Moreover, this ideal can be assumed to be zero, for otherwise our proposition would follow trivially.

Thus we suppose without loss of generality that $\mathfrak{S} \cap [\mathfrak{R}, \mathfrak{R}] = (0)$. Here $\mathfrak{S}$ is abelian. Furthermore, S^* clearly contains a regular element of R^* and thus $\mathfrak{S}$ contains a regular element X of $\mathfrak{R}$ (cf. [11]). Let $\mathfrak{A}$ be the subset of $\mathfrak{R}$ which is annihilated by some power of $\operatorname{ad} X$. As is well-known, $\mathfrak{A}$ is a Cartan subalgebra of $\mathfrak{G}$, and as such, is its own normalizer. It is proved in [11] that $\mathfrak{G} = \mathfrak{A} + \mathfrak{R}^\infty$ with $\mathfrak{A} \cap \mathfrak{R}^\infty \subset [\mathfrak{R}^\infty, \mathfrak{R}^\infty]$. Hence we have here $\mathfrak{G} = \mathfrak{A} + \mathfrak{R}^\infty$ (semi-direct) with $\mathfrak{A} \supset \mathfrak{S}$.

Let A^* be the subgroup of G^* which corresponds to $\mathfrak{A}$. Then clearly A^* is the connected component of the identity of its normalizer $N(A^*)$. We assert now that A^* coincides with $N(A^*)$. To prove this we need only show that $N(A^*) \cap R^{*\infty}$ is connected, since $R^* = A^* \cdot R^{*\infty}$. Now let x be in $N(A^*) \cap R^{*\infty}$, and let X be an element of the Lie subalgebra R such that $\exp X = x$. Denoting by $\alpha(x)$ the restriction to $\mathfrak{R}$ of $\operatorname{ad} X$, we have that $\alpha(X)$ is nilpotent and $\exp \alpha(X)$ keeps $\mathfrak{A}$ invariant. Hence $\alpha(X) = \log \exp \alpha(X)$ keeps A invariant and thus $\alpha(\exp tX) = \exp t\alpha(X)$ keeps $\mathfrak{A}$ invariant for every t and $N(A^*)$ contains a connected subgroup through x. From this it follows that $N(A^*) \cap R^*$ is connected as well as discrete, and is thus {identity}. It follows that $N(A^*) = A^*$.

Knowing this, we assert that $D_R{}^* \subset A^*$. For from the definition of $\mathfrak{A}$, it is seen to be the centralizer of $\mathfrak{C} \cap \mathfrak{R}$ in $\mathfrak{R}$ and it is hence invariant under $\operatorname{Ad} L^*$. Hence $\operatorname{Ad} D_R{}^*(\mathfrak{A})$ $\operatorname{Ad} D^* \cdot \operatorname{Ad} D_L(\mathfrak{A}) \subset \mathfrak{A}$ and thus $D_R{}^* \subset N(A^*) = A^*$.

Inasmuch as $G^* = L^* \cdot A^* \cdot R^{*\infty}$ topologically and $D^* \subset D_L{}^* D_R{}^* \subset L^* A^*$, we conclude that $G = (LA) \cdot R^\infty$ (semi-direct), where $L = f(L^*)$ and $A = f(A^*)$, $C \subset L \cdot A$ and, in addition, R^∞ is an abelian simply connected normal analytic subgroup that is non-trivial. Hence by hypotheses, $G = \mathcal{C}\ell\, C R^\infty$. But the foregoing topological fibering of G yields $\mathcal{C}\ell\, C R^\infty = (\mathcal{C}\ell\, C) \cdot R^\infty = C \cdot R^\infty$. Hence $G = CR^\infty$, and Proposition 2.1 holds with $W = R^\infty$. Proof of Proposition 2.1 is now complete.

Note. The subgroup W of Proposition 2.1 can be taken to be non-trivial. For if W were trivial, we would have $G = C$. Since G has a non-trivial radical R, it contains a non-trivial normal abelian analytic subgroup W', namely, the last non-trivial subgroup in the sequence of successive commutator subgroups of R. Replacing W by W', our assertion follows.

THEOREM 2.1. *Let G be a connected Lie group, and C a closed connected subgroup which is self-adjoint in G modulo the radical. Let M_C be a* *maximal compact subgroup of C and let M be a maximal compact subgroup of G which includes M_C. There exist M_C-invariant exp-sets F and E such that $C = M_C \cdot E$ and $G = M \cdot F \cdot E$ topologically.*

Proof is by induction on the dimension of G. We consider three cases, the third one being the case of trivial radical.

Case 1. G contains a non-trivial central toroid subgroup T.

By the induction hypothesis $G^* = G/T = M^* \cdot F^* \cdot E^*$ topologically where F^* and E^* are M_C^* $(= M_C T/T)$-invariant exp-sets, $C^* = CT/T = M_C^* \cdot E^*$, and $M^* = MT/T$. We know that F^* arises out of independent M_C^*-invariant linear subspaces $\mathfrak{F}_1^*, \cdots, \mathfrak{F}_p^*$ and that E^* arises similarly out of $E_1^*, \cdots, E_q^*$. Let ϕ denote the homeomorphism of G onto G^*, and let $\mathfrak{F}_i = d\phi^{-1}(\mathfrak{F}_i^*)$, $\mathfrak{E}_j = d\phi^{-1}(\mathfrak{E}_j^*)$ for each i, j. Clearly, each $\mathfrak{F}_i$, $\mathfrak{E}_j$, is invariant under the compact linear group $\mathrm{Ad}\, M_C$. Let $\mathfrak{F}_i$ be an $\mathrm{Ad}\, M_C$ invariant complement to $\mathfrak{T}$ in $\mathfrak{F}_i$, and let $\mathfrak{E}_j$ be an $\mathrm{Ad}\, M_C$ invariant complement to $\mathfrak{T}$ in $\mathfrak{E}_j$ which is also included in $\mathfrak{C} \cap \mathfrak{E}_j$ for each i and j. That $\mathfrak{E}_j$ can be so chosen follows from the fact that $\mathfrak{C} \cap \mathfrak{E}_j$ is $\mathrm{Ad}\, M_C$ invariant and spans $\mathfrak{E}_j$ together with $\mathfrak{T}$. Clearly $d\phi$ is bi-unique on each $\mathfrak{F}_i^*$ and $\mathfrak{E}_j^*$. Set $F = \exp \mathfrak{F}_1 \cdots \exp \mathfrak{F}_p$ and $E = \exp \mathfrak{E}_1 \cdots \exp \mathfrak{E}_q$. From $G^* = (MFE)^* = \phi(MFE)$, we deduce $G = MFET = MTFE = MFE$. To prove that $G = M \cdot F \cdot E$ topologically, we must prove that the mapping $\theta \colon (m, f, e) \rightarrow mfe$ of $M \times F \times E$ onto G is a homeomorphism.

θ is one-to-one. For knowing that $G^* = M^* \cdot F^* \cdot E^*$ topologically, it is sufficient to prove that ϕ is bi-unique on F and E.

Knowing that $F^* = \exp \mathfrak{F}_1^* \cdots \exp \mathfrak{F}_p^*$ topologically, it is sufficient in order to prove that ϕ is bi-unique on each $\exp \mathfrak{F}_i$.

Suppose therefore that $f_i = \exp X_i$ $(i = 1, 2)$ with X_i in $\mathfrak{F}_h$ and $\phi(f_1) = \phi(f_2)$. Then $\exp d\phi(X_1) = \phi(\exp X_1) = \phi(\exp X_2) = \exp d\phi(X_2)$ where $d\phi$ is the differential of ϕ at the identity. Since $d\phi$ is one-to-one on $\mathfrak{F}_h$ by choice of $\mathfrak{F}_h$, and exp is one-to-one on $d\phi(\mathfrak{F}_h) = \mathfrak{F}_h^*$, we conclude $X_1 = X_2$ and $f_1 = f_2$. Similarly ϕ is bi-unique on E.

It follows as point out above that the mapping θ is one-to-one. That θ is open as well as continuous follows from the invariance of domain theorem for Euclidean space applied to the locally Euclidean space $M \times F \times E$ whose dimension is the same as the dimension of G. Thus θ is a homeomorphism and $G = M \cdot F \cdot E$ topologically. Since $\phi(C) \subset M^* \cdot E^*$ and $T \subset M$, we have $C \subset ME$ and hence $C = (C \cap M) \cdot E$ topologically. Since $C \cap M = M_C$, we have $C = M_C \cdot E$ topologically.

Case 2. G has no non-trivial central toroidal subgroup but has a non-trivial radical.

Case 2a. There exists a non-trivial abelian analytic normal V with CV closed.

Let R denote the radica of G and let N denote the analytic subgroup corresponding to the maximum nilpotent ideal of R. By Lemma 2.2, N is a closed simply connected subgroup of G. V is included in N. Since N is simply connected, the analytic subgroup V is simply connected (cf. [4]) and therefore V is a non-trivial vector group. We now apply the induction hypothesis to $G^* = G/V = M^* \cdot F^* \cdot E^*$ where F^* and E^* are $M_C{}^*$-invariant exp-sets, and $C^* = CV/V = M_C{}^* \cdot E^*$ topologically, M_C being $M \cap C$, and M^* denoting MV/V.

Let ϕ denote the homomorphism of G onto G^*. Clearly $\phi(M_C) \subset M_C{}^*$. Following the procedure used in Case 1 we can define M_C-invariant exp-sets F' and E' which map bi-uniquely onto F^* and E^*. Let $\mathfrak{V}_1$ denote an Ad M_C-invariant subspace of the abelian Lie algebra $\mathfrak{V}$ which is a complement to $\mathfrak{C} \cap \mathfrak{V}$ in $\mathfrak{V}$, and let V_1 denote the corresponding vector subgroup of V. Set $F = F'V_1$ and $E = (C \cap V)^c \cdot E'$, where $(C \cap V)^c$ is the connected component of the identity in $C \cap V$. Then from $G^* = \phi(MF'E)$ we deduce

$$G = MF'EV = MF'VE = MF'V_1(C \cap V)^c \cdot E = MFE.$$

To prove that $G = M \cdot F \cdot E$ topologically, it suffices to prove that

$$G = M \cdot F' \cdot V_1 \cdot (C \cap V_1)^c \cdot E'$$

topologically. Inasmuch as $V = V_1 \cdot (C \cap V)^c$ topologically, it suffices to prove $G = M \cdot F' \cdot V \cdot E'$ topologically. Let θ denote the mapping $(m, f, v, e) \rightarrow mfve$ of $M \times F' \times V \times E'$ onto G.

θ is bi-unique. For $m_1 f_1 v_1 e_1 = m_2 f_2 v_2 e_2$ (m_i in M, etc,) implies $\phi(m_1)\phi(f_1)\phi(e_1) = \phi(m_2)\phi(f_2)\phi(e_2)$. Since $M \cap V$ is a compact subgroup of V, it consists only of the identity and hence ϕ is one-to-one on M. By construction ϕ is one-to-one on F' and E'. Hence $G^* = \phi(M) \cdot \phi(F') \cdot \phi(E')$ topologically implies $\phi(m_1) = \phi(m_2)$, $\phi(f_1) = \phi(f_2)$, $\phi(e_1) = \phi(e_2)$ and consequently, $m_1 = m_2$, $f_1 = f_2$, $e_1 = e_2$. It follows that $v_1 = v_2$ and θ is bi-unique. That θ is a homeomorphism follows as in Case 1.

Case 2b. There exists a non-trivial normal analytic subgroup V with $\mathcal{C}\ell\, CV \neq G$.

Take H to be the subgroup $\mathcal{C}\ell\, CR$ if this does not coincide with G, otherwise take it to be $\mathcal{C}\ell CV$. Then H, being a closed connected subgroup, is analytic and moreover, its dimension is smaller than the dimension of G.

We assert that C and H are self-adjoint modulo the radical in H and G respectively. For by hypothesis there is a star automorphism of G/R which keeps invariant CR/R and this clearly keeps invariant $\mathcal{C}\ell(CR/R) = (\mathcal{C}\ell CR)/R$. Thus the assertion is true for H if $G \neq \mathcal{C}\ell CR$. If, on the other hand, $G = \mathcal{C}\ell CR$, then by a result of Malcev (cf. first paragraph of proof of Proposition 2.1), $[G, G] \subset CR$ and thus

$$G/R = [G/R, G/R] = [G, G]R/R = CR \neq R$$

so that $\mathcal{C}\ell CV$ is self-adjoint modulo the radical in G in this case. Thus in any event, H is self-adjoint modulo the radical in G.

As for C being self-adjoint modulo the radical in H, the star automorphism of G which keeps H invariant induces a star automorphism of H/S which keeps invariant CS/S, S denoting the radical of H. Thus C is self-adjoint modulo the radical in H.

Applying the induction assumption to the pair C, H, we obtain $C = M_C \cdot E$ and $H = M_H \cdot F_1 \cdot E$ topologically, where M_H is a maximal compact subgroup of G containing M_C, E, and F_1 and hence $F_1 \cdot E$ are M_C-invariant exp-sets.

We now apply Case 1 to the pair H, G, and conclude that $H = M_H \cdot E_1$, $G = M \cdot F_2 \cdot E_1$ where E_1 and F_2 are M_H-invariant exp-sets and M is a maximal compact subgroup of G containing M_H.

Any element y of E can be written uniquely as $y = m_1 f_1 e$ with m_1, f_1, e in M_H, F_1, E respectively and depending continuously on y. Hence any element x of G can be expressed uniquely as $x = mf_2 \cdot m_1 f_1 e = mm_1 \cdot m_1^{-1} f_2 m_1 \cdot f_1 \cdot e_1$ with the factors in M, F_2, F_1, E respectively and depending continuously on x. Hence $G = M \cdot F_2 \cdot F_1 \cdot E_1 = M \cdot F \cdot E_1$ topologically where F is the M_C-invariant exp-set $F_2 F_1$. Proof is now complete in this case.

Case 2c. $R \neq 0$ and for any non-trivial normal abelian analytic subgroup V, $G = \mathcal{C}\ell CV$.

In this case, there exists a normal analytic abelian non-trivial subgroup W such that $G = CW$ by Proposition 2.1 and the Note following it. Thus we may apply the result in Case 2a to obtain the theorem.

The induction hypothesis has thus been proved to hold for G in case G has a non-trivial radical. It remains only to consider

Case 3. G has no radical.

Here G is semi-simple and we can apply Theorem 5 of Section 1 in [10] which asserts: Let $\mathfrak{K} + \mathfrak{E}$ be a Cartan decomposition for the real semi-simple Lie algebra $\mathfrak{G}$ and let $\mathfrak{E}'$ be a linear subspace of $\mathfrak{E}$ such that $[X, [X, Y]]$ is in E' for any X, Y in E'. Let $\mathfrak{F}$ be the orthogonal complement to $\mathfrak{E}'$ in $\mathfrak{E}$ with respect to the Killing form, let $F = \exp \mathfrak{F}'$, $E' = \exp \mathfrak{E}'$. Then $G = K \cdot F \cdot E'$ topologically where K is the analytic subgroup determined by $\mathfrak{K}$.

In Case 3, C is self-adjoint in G and therefore there exists a Cartan decomposition $\mathfrak{K} + \mathfrak{E}$ with $\mathfrak{C} = \mathfrak{C} \cap \mathfrak{K} + \mathfrak{C} \cap \mathfrak{E}$. Set $\mathfrak{E}' = \mathfrak{C} \cap \mathfrak{E}$. Then $\mathfrak{E}'$ satisfies the hypothesis of the cited theorem and therefore $G = K \cdot F \cdot E'$ topologically, notation being as above. In addition, we have $C = (C \cap K) \cdot E'$ topologically. For the subsets $C \cap K$ and E' being closed, $(C \cap K) \cdot E'$ is a closed subset of C. On the other hand, the mapping $\theta \colon (c, e) \to ce$ of $(C \cap K) \times E'$ into C is one-to-one and continuous; since $(C \cap K) \times E'$ is locally Euclidean and of the same dimension as C, the mapping θ is open by the theorem on the invariance of domain. As a result, $(C \cap K) \cdot E'$ is open and closed in the connected group C and hence coincides with C. Inasmuch as $\mathfrak{E}$ is invariant under $\operatorname{ad} \mathfrak{K}$ we conclude that E' and F' are $(C \cap K)$-invariant exp-sets.

The choice of F and E is not yet complete inasmuch as K need not be compact. We know only that the image of K in the adjoint group of G is compact. This implies that the adjoint groups of K and $C \cap K$ are compact. It follows at once that each is the direct product of its maximum compact subgroup and a vector group (cf. [9], Lemma 2.8). Let M_1 be the maximum compact subgroup of C and let W be a vector subgroup such that $C = M_1 \cdot W_1$ (direct). Let W_2 be a vector subgroup of K such that $K = (MC) \cdot W_2$ (direct) where M is the maximum compact subgroup of K. Set $E = W_1 E'$ and $F = W_2 F'$. Since $W_1 \subset C \cap K$, $W_2 \cdot F' = F' \cdot W_2$, we get

$$\begin{aligned} G &= M \cdot W_1 \cdot W_2 \cdot F' \cdot E' = M \cdot W_2 \cdot W_1 \cdot F' \cdot E' \\ &= M \cdot W_2 \cdot F' \cdot W_1 \cdot E' = M \cdot F \cdot E \text{ topologically.} \end{aligned}$$

Finally, M_1 is a maximum compact subgroup of C. For let M_C be a maximal compact subgroup of C which contains M_1. Since $C = M_1 \cdot E$ topologically, G/M_1 is homeomorphic to Euclidean space and it is fibered by the compact fibers M_C/M_1 with $G/M_1/M_C/M_1 = G/M_C$ as base space. Since Euclidean space cannot be fibered by compact fibers containing more than one point ([2] or [12]), we conclude that $M_1 = M_C$. Proof of the theorem is now complete.

Note. The theorem from [10] that was cited in Case 3 arises out of the following theorem, which the author proved with precisely the above application in mind.

Let $\mathfrak{E}$ be any real linear subspace of $\mathfrak{S}$, the set of real symmetric matrices, and let $\mathfrak{F}$ be an orthogonal complement to $\mathfrak{E}$ in $\mathfrak{S}$ with respect to the bilinear form $\operatorname{Tr} XY$. Then any positive definite symmetric matrix can be written uniquely and continuously as efe with e in $\exp \mathfrak{E}$ and f in $\exp \mathfrak{F}$, if and only if $[\mathfrak{E}, [\mathfrak{E}, \mathfrak{E}]] \subset \mathfrak{E}$.

When we specialize the choice of C to a maximal compact subgroup M, we obtain the following sharpened form of the theorem on the topological decomposition of a Lie group.

THEOREM 2.2. *Let G be a connected Lie group and let M be a maximal compact subgroup of G. Then $G = M \cdot U$ topologically, where U is an M-invariant exp-set.*

In particular, U is a Euclidean subspace with $mUm^{-1} = U$ for all m in M, and M operating on U by inner automorphisms is equivalent to the linear group $\operatorname{Ad} M$ operating on a linear subspace of G.

Theorem 2.2 allows us to sharpen the theorem on the conjugacy of maximal compact subgroups of connected Lie groups.

Remark. That any two maximal compact subgroups of a connected Lie group G are conjugate under inner automorphisms has been proved by Malcev ([7]) and Iwasawa ([6]). Each proof takes as starting point the theorem of E. Cartan that maximal compact subgroups of a semi-simple group are conjugate (cf. [3] or [10]). It is of interest to observe that the idea introduced by Cartan to handle the semi-simple case can be used to handle the general case. We sketch the argument. Proceeding by induction in the case that G is not semi-simple, we reduce to the case that G contain a non-trivial closed normal vector subgroup V (cf. proof of Theorem 2.1).

Applying the induction assumption to G/V, it follows that if M_1 and M_2 are maximal compact subgroups, then M_2 is conjugate under an inner automorphism to a subgroup of $M_1 V$. Hence we lose no generality in assuming that $M_2 \subset M_1 V = G$, with G connected. M being compact, $M_1 \cap V = (\text{identity})$ so that $G = M_1 \cdot V$ topologically. Hence $G/M_1 = V$. Now upon selecting logarithmic coordinates in V, it is seen that G operating on V is equivalent to group of affine transformations, the elements of V operating as translations. Hence any compact subgroup of G admits a fixed point in V, the centroid of any of its orbits. Thus there is an element x in G such that $M_2 x M_1$ is simply the coset xM_1; that is, $M_2 x M_1 \subset xM_1$ and $x^{-1} M_2 x \subset M_1$. Thus M_1 and M_2 are conjugate. That M_1 is connected follows from the connectedness of G.

THEOREM 2.3′. *Let M_1 and M_2 be any two maximal compact subgroups* *of the connected Lie group G. Then there is an element x in G such that $xM_1x^{-1}=M_2$ and $xmx^{-1}=m$ for all m in $M_1 \cap M_2$.*

Proof. We begin with the fact that any two maximal compact subgroups of G are conjugate under an inner automorphism. Thus we may assume that M_1 is the subgroup M of Theorem 2.2, i.e., there is a subspace U such that $G=M_1 \cdot U$ topologically and $mUm^{-1}=U$ for all m in M. We know there is an element y in G such that $yM_1y^{-1}=M_2$. Now y can be written as xn with n in M, and x in U. Then $xM_1x^{-1}=xnM_1n^{-1}x^{-1}=yM_1y^{-1}=M_2$. Thus $xM_1x^{-1}=M_2$ and hence $x(M_1 \cap M_2)x^{-1} \subset M_1$. For any element m in $M_1 \cap M_2$ we have the relation $xmx^{-1} \cdot x = xm = m \cdot m^{-1}xm$.

Since $xmx^{-1} \varepsilon M_1$, $m^{-1} \varepsilon U$, and $G=M_1U$ topologically, we infer $xmx^{-1}=m$, which proves the theorem.

More generally, we can assert

THEOREM 2.3. *Let N_1 and N_2 be two compact subgroups of the connected Lie group G which are conjugate under an inner automorphism θ_1. Let M_1 and M_2 be any maximal compact subgroups which include N_1 and N_2 respectively. Then there exists an inner automorphism θ of G which sends M_1 onto M_2 and which coincides with θ_1 on N_1.*

Proof. Let $M_3=\theta_1(M_1)$. Applying Theorem 2.3′ to M_3 and M_2, we find an inner automorphism θ_2 sending M_3 into M_2 and keeping $M_3 \cap M_2$ pointwise fixed. The inner automorphism $\theta=\theta_2\theta_1$ has the desired properties.

3. Retraction of Klein spaces.

Throughout this section, G denotes a connected closed subgroup, and M any maximal compact subgroup of G which includes a maximal compact subgroup of C.

THEOREM 3.1. *G/C can be retracted by a strong deformation retraction to a subset homeomorphic to the h-compact space $M/M \cap C$. Thus G/C and $M/M \cap C$ have the same homotopy type.*

Proof. For convenience, let M_C denote $M \cap C$. By hypothesis M_C is a maximal compact subgroup of C. By Theorem 2.2, C/M_C is topologicaly a Euclidean space and is thus solid (cf. [13] for definition). Moreover, G/M_C is a fiber bundle with G/C as base space and C/M_C as fiber, the projection p being $[xM_C] \rightarrow [xC]$, (x in G).

Let ϕ denote the mapping $[xC] \rightarrow [xM_C]$ for x in M. Since $M \cap C = M_C$, ϕ is well-defined and is a continuous cross-section of the subset MC/C of G/C into the bundle G/M_C. Since the fiber C/M_C is solid, ϕ can be extended to a continuous cross-section of G/C into G/M_C ([13]). We denote this extension by ϕ also. From the fact that $G = U \cdot M$ topologically with U a Euclidean subspace, it follows at once that there is a strong deformation retraction $S_t (0 \leqq t \leqq 1)$ of G/M_C onto the subset M/M_C (the points of M/M_C being fixed under the deformation). We define the retraction R_t of $G_C C$ onto itself as $R_t = p S_t \phi (0 \leqq t \leqq 1)$. Then for any x in M,

$$R_t([xC]) = pS_t([xM_C]) = p([xM_C]) = [xC].$$

Moreover, $R_0 =$ identity, and $R_1(G/C) = MC/C = M/M \cap C$. Thus R_t is the desired strong deformation retraction of G/C.

COROLLARY 1. *If S is a closed subgroup of G and G/S is simply connected, then the conclusion of the above theorem holds.*

Proof. If G/S is simply connected, then S is connected.

COROLLARY 2. (Montgomery, [8]). *If G is a Lie group that is transitive on a simply connected compact space, then there is a compact subgroup of G that is transitive on the space. Thus every simply connected compact Klein space is h-compact.*

Proof. We may assume that our space is G/C where C is connected, and G/C is a compact manifold. Since a compact manifold cannot be deformed to a proper subset, $G/C = MC/M$, i.e., M is transitive on G/C and $G/C = M/M \cap C$.

4. Covariant fiberings. Let B be a Lie group, B' a closed subgroup of B, let F denote a topological space, and let λ denote an anti-homomorphism of B' into the group of all homeomorphisms of F onto itself (cf. [0]). Assume that B operates on F as a topological transformation group in the usual sense, i.e., the mapping $(g, f) \rightarrow \lambda(g)f$ of $B' \times F$ into F is continuous. We define the subset $[b, f]$ for any point (b, f) in $B \times F$ as $\{(bg, \lambda(g)f) \mid \text{all } g \varepsilon B'\}$. If U is a subset of $B \times F$, we denote by $[U]$ the set $\{[b, f] \mid (b, f) \varepsilon U\}$.

Definition. (B, F, B', λ) denotes the topological space whose points are the subsets $[b, f]$ *of* $B \times F$, a neighborhood of $[b, f]$ being defined as $[U]$ where U is any neighborhood of (b, f) *in* $B \times F$.

Let π denote the mapping $(b, f) \rightarrow [b, f]$ of $B \times F$ onto (B, F, B', λ). Then by definition essentially, π is continuous and open. (B, F, B', λ) can be thought of as the space obtained from $B \times F$ by taking as points the (non-singular) orbits of the transformation group $R(B')$, where $R(g)$ is the mapping $(b, f) \rightarrow (bg, \lambda(g)f)$, $b \varepsilon B$, $g \varepsilon B'$, $f \varepsilon F$.

We next define on (B, F, B', λ) a fiber bundle structure, with base space B/B', fiber F, and as structure group the Lie group $\lambda(B')$. Namely, let p be the mapping $[b, f] \rightarrow bB'$ of (B, F, B', λ) onto B/B'. Let $U_1, \cdots, U_i, \cdots$ be a covering of B/B' by neighborhoods such that there exists a continuous cross-section ψ_i of U_i into the bundle B fibered by cosets B' (b in B). Define $\phi_i : p^{-1}(U_i) \rightarrow U_i \times F$ by the relation $\phi_i(\psi_i(x), f) = (x, f)$, ($x$ in U_i, f in F). It is easily seen that ϕ_i constitutes a set of coordinate functions which defines (B, F, B', λ) as a bundle with fiber F and structure group $\lambda(B')$. We shall employ the symbol (B, F, B', λ) to denote the above bundle structure as well as the underlying space, when there is no danger of confusion. It is clear from the choice of coordinate functions that the associate principal bundle of (B, F, B', λ) is the bundle B fibered by cosets bB'.

Definition. Suppose B and F are closed subsets of a prescribed Lie group G. Suppose moreover that F is invariant under inner automorphisms by elements of B', that F contains the identity element of G, and that $\lambda(g)$ denotes the transformation $f \rightarrow g^{-1}fg$ where $f \varepsilon F$ and $g \varepsilon B'$. We then denote (B, F, B', λ) by the symbol $(B \times F)_{B'}$.

Let (B, F, B', λ) be given. If the group of transformations $\lambda(B)$ has a fixed point p_0 in F, then the mapping $[bB'] \rightarrow [b, p_0]$ of B/B' into (B, F, B', λ) is well defined and is in fact a continuous cross-section of the base space B/B' into the fiber bundle (B, F, B', λ). In particular, the mapping $[bB'] \rightarrow [b, e]$ of B/B' into $(B \times F)_{B'}$ is a continuous cross-section, where e is the identity of the prescribed group G.

LEMMA 4.1. *Let G be a Lie group, let C and B be closed subgroups, and let F, E be two subspaces of G such that, $G = B \cdot F \cdot E$ (topologically). Assume*

(1) *$C = B' \cdot E$ (topologically) where B' is a subgroup of B.*

(2) *$g^{-1}Fg \ F$ for all $g \varepsilon B'$.*

(3) *F contains the identity element of G.*

Then $G/C = (B \times F)_{B'}$ under the homeomorphism $[bfC] \rightarrow [b, f]$.

Proof. Let θ denote the mapping $(b, f) \rightarrow bfC$ of $B \times F$ onto G/C. For each point of G/C, $\theta^{-1}[bfC]$ is the totality of pairs (b^*, f^*) in $B \times F$ such that $b^*f^* = bfC = bfge = bg \cdot g^{-1}fg \cdot e$ with $c \varepsilon C$, $g \varepsilon B'$, $e \varepsilon F$. Hence $b^* = bg$, $f = g^{-1}fg$, and so $\theta^{-1}[bfC]$ is the subset $[b, f] \subset B \times F$. As a result, the correspondence $\pi\theta^{-1} : \theta(b, f) \rightarrow (b, f)$ is a well defined one-to-one correspondence between G/C and $(B \times F)_{B'}$. Since both θ and π are open and continuous mappings, $\pi\theta^{-1}$ is a homeomorphism. Proof of the lemma is now complete.

Remark 4.1. Suppose that $G = B \cdot F \cdot E$ topologically and $C = B' \cdot E$ topologically, with B' a subgroup of B as above. Assume moreover that $bFb^{-1} \subset F$ for all b in B. Then G is also $F \cdot B \cdot E$ topologically. From these two decompositions of G we obtain two different fiber bundles (in general) which are associated with the *same* fibering of G/C into the subsets bF $(= Fb)$ with b in B; the bundles are (B, B', F, λ_1) and (B, B', F, λ_2) respectively, where λ_1 corresponds to the operation of B' on F by inner automorphisms, and $\lambda_2(B')$ is the identity transformation. The bundle $(B, B', F, \lambda_1) = (B \times F)_{B'}$ has as structure group $\lambda_1(B')$ and the other has as structure group $\lambda_2(B)$ $=$ (identity), these groups being different in general. Since these correspond to the same fibering of G/C, they are equivalent when the structure group of each bundle is suitably augmented.

Remark 4.2. The group B operates on (B, B', F, λ) by the operations $L_b{}^*: [xf] \to [bx, f]$, where b, x are in B and f is in F. $L_b{}^*$ is a well defined bundle map of (B, B', F, λ) onto itself, and $L_B{}^*$ permutes the fibers transitively.

In the special case that $(B, B', F, \lambda) = (B \times F)_{B'}$, and the subsets B and F of G are as in Lemma 4.1, the operation $L_B{}^*$ is the operation of B on G/C by left translation. In this case, the operation $R(b): (x, f) \to (xb, b^{-1}xb)$ with b in B' corresponds to the operation of b on BF by right translation.

Definition. Let G be a connected Lie group which operates transitively on the space H. A *G-covariant fibering* of the differentiable manifold H is a differentiable fibering into Euclidean fibers (i. e., the fibers are submanifolds and each fiber has as neighborhood an open set of fibers with a differentiable submanifold as cross-section set) such that for some maximal compact subgroup M of G,

1) M permutes the fibers transitively.

2) For some fiber F, the subgroup M_F of M which keeps F invariant is equivalent to a group of linear transformations on the Euclidean space F.

Condition 2) is obviously equivalent to the condition obtained on replacing "some fiber F" by "each fiber F," in view of the fact that M permutes the fibers transitively.

THEOREM 4.1. *Let G be a connected Lie group which operates transitively on the space H with a connected isotropy subgroup C. Assume that C is self-adjoint modulo the radical. Then H admits a G-covariant fibering.*

Proof. By Theorem 2.1, $G = M \cdot F \cdot E$ topologically, where M is a maximal compact subgroup, $C = M_C \cdot E$ topologically, with $M_C = M \cap C$, and

F is an M_C-invariant exp-set. By Lemma 4.1, $H = G/C$ is the underlying space of the fiber bundle $(M \times F)_{M_C}$. Since M permutes the fibers of this bundle transitively (Remark 4.2), F is Euclidean. To conclude the proof of our theorem it suffices to verify that a) M_F, the subgroup of M keeping the fiber F invariant is equivalent to a linear group, and b) $(M \times F)_{M_C}$ is a differentiable fibering of H. Part a) follows from the observation that $M_F = M_C$ and M_C operating on F is equivalent to the linear group $\mathrm{Ad}\, M_C$ operating on a linear subspace of G by definition of F. Part b) follows from the fact that decomposition $G = M \cdot F \cdot E$ is differentiable around the identity element at least, as can be verified without difficulty.

COROLLARY 1. *If C is a closed connected semi-simple subgroup of the connected Lie group G, then G/C admits a G-covariant fibering.*

Proof. Let R be the radical of G. Then CR/R is a semi-simple subgroup of the semi-simple group G/R and is self-adjoint by Theorem 6 of [10].

COROLLARY 2. *If C is a closed connected subgroup of the connected solvable Lie group G, then G/C admits a G-covariant fibering.*

Proof. G is its radical and CG/G is self-adjoint in G/G.

COROLLARY 3. *If C is a compact connected subgroup, then G/C admits a G-covariant direct product fibering.*

Proof. This follows directly from Theorem 2.2 applied to a maximal compact subgroup containing C. (cf. Remark 4.1).

Remark 4.3. The decomposition of a semi-simple Lie group G without center into $K \cdot S$ where K is a maximal compact subgroup and S is a subgroup homeomorphic to Euclidean space (due to Iwasawa [6]) allows one to conclude that G/K has a covariant fibering but does *not* allow this conclusion for G/K' when $K' \subset K$, inasmuch as S is not invariant under inner automorphisms from K'.

5. Examples.

We will now consider some cases where the fiberings of Section 4 turn out to be direct product fiberings when the group of the bundle is suitably enlarged.

Definition. Let F be a locally compact Hausdorff space. By $\mathrm{aut}\, F$ is meant the group of homeomorphisms of F onto itself topologized so as to be a topological transformation group operating on F. (cf. [0], Arens).

In the case that F is Euclidean space, the compact open topology satisfies the above condition.

Conditions that the fibering of Section 4 be product fiberings can be stated most simply in terms of mappings into $\operatorname{aut} F$.

We consider first the fibering defined for the space $(B \times F)_{B'}$ and employ the notation of Section 4. In particular, $\lambda(g)$ denotes the homeomorphism $f \rightarrow g^{-1}fg$ for $f \varepsilon F$, $g \varepsilon B'$. Since $\lambda(g_1 g_2) = \lambda(g_2)\lambda(g_1)$, λ is an anti-homomorphism of B' into $\operatorname{aut} F$.

THEOREM 5.1. *The fibering of $(B \times F)_{B'}$ is a direct product decomposition if and only if there is a continuous mapping $u\colon B \rightarrow \operatorname{aut} F$ such that $u(x) = u(xg)\lambda(g)$ for each g in B'.*

Proof. If the fibering is a direct product decomposition, then there exists a fiber-preserving homeomorphism ϕ of $(B/B) \times F$ onto $(B \times F)_{B'}$ such that $\phi([xB'],{}_1) = (x,f)\,(=(xg, g^{-1}fg) = (xg, \lambda(g)f)$ for g in B'). Define $u(x)$ to be the mapping $f \rightarrow f_1$ of F onto F, where $x \varepsilon B$. Since ϕ is a fiber preserving homeomorphism, $u(x)$ is a homeomorphism. Moreover u is a continuous mapping of B into $\operatorname{aut} F$. Finally, since (x,f) and $(xg, \lambda(g)f)$, g in B', represent the same point of $(B \times F)_{B'}$, we have the identity $u(x)(f) = u(xg)(\lambda(g)f)$ for any f in F, i.e., $u(x) = u(xg)\lambda(g)$ for all g in B'.

Conversely, suppose there is such a mapping u. Then consider the mapping $\theta\colon (x,f) \rightarrow ([xB'], u(x)f)$ of $B \times F$ into $(B/B') \times F$. The mapping θ is single-valued, open and continuous. Thus θ is a fiber preserving homeomorphism of $(B \times F)_{B'}$ with $(B/B') \times F$.

Let $\mathcal{R}$ and $\mathcal{C}$ denote the field of real complex numbers respectively.

We employ the notation $Sl(n, \mathcal{R})$, $Sl(n, \mathcal{C})$, $SO(n, \mathcal{R})$, $SO(n, \mathcal{C})$, $SU(n)$ to denote the linear groups over real and complex numbers of determinant 1, the orthogonal groups over $\mathcal{R}$ and $\mathcal{C}$ of determinant 1, and the unitary group of determinant 1.

Example 1. We examine the fibering of the space G/C, where G is the group of matrices $Sl(n+1, \mathcal{R})$ and C is the subgroup $Sl(n, \mathcal{R})$ of the form $\begin{pmatrix} 1 & 0 \\ 0 & g \end{pmatrix}$, g being an n by n matrix. Let $\mathfrak{K}$ and $\mathfrak{S}$ denote the totality of skew-symmetric and symmetric matrices, respectively. Let $\mathfrak{E} = \mathfrak{C} \cap \mathfrak{S}$ and define $\mathfrak{F} = \{X \mid X \varepsilon \mathfrak{S},\ \operatorname{Tr} XY = 0 \text{ for all } Y \varepsilon E\}$. Then $\mathfrak{F}$ is the totality of symmetric matrices (s_{ij}) with $s_{ij} = 0$ if $i > 1$ and $j > 1$. We set $\mathfrak{K}' = \mathfrak{C} \cap \mathfrak{K}$, $\mathfrak{F} = \exp \mathfrak{F}$. We define $\lambda(g)$ to be the mapping $f \rightarrow g^{-1}fg$ in $\operatorname{aut} F$, where g is in the analytic subgroup K'. Let $A(g)$ denote the restriction of $\operatorname{Ad} g$ to $\mathfrak{F}$ where $g \varepsilon K'$. Let log denote the homeomorphism $\exp X \rightarrow X$ of F onto $\mathfrak{F}$. Since $\lambda(g)(\exp X) = \exp(\operatorname{Ad} g^{-1}(X))$ for all $X \varepsilon \mathfrak{F}$, $g \varepsilon K'$, we have

$\log\cdot\lambda(g)=A(g^{-1})\cdot\log$, that is, the homeomorphism log is equivalent with respect to the isomorphism $\lambda(g)\rightarrow A(g^{-1})$. We now look at the transformation group $A(K')$. If $X\,\varepsilon\,\mathfrak{F}$, then

$$X=\begin{pmatrix} a & b' \\ b & c \end{pmatrix}, \qquad c=-a/n\cdot(\text{identity})$$

and

$$A(g)X=\begin{pmatrix} 1 & 0 \\ 0 & g \end{pmatrix}\begin{pmatrix} a & b' \\ b & c \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & g \end{pmatrix}^{-1}=\begin{pmatrix} a & b'g^{-1} \\ gb & c \end{pmatrix}, g\,\varepsilon\,K'.$$

This identity has the following interpretation. Let ϵ_i denote the $(n+1)$-tuple with components δ_{ij} $(i,j=1,\cdots,n+1)$. Let X_i denote the matrix of F whose p,q-th coefficient is $\max\limits_{i}\{\delta_{ip}\delta_{iq},\delta_{iq}\delta_{ip}\}$. Let ϕ denote the unique linear mapping of F onto the Cartesian space $\mathfrak{R}^{n+1}$ (on which $G=Sl(n+1,\mathfrak{R})$ operates) such that $\phi(X_i)=\epsilon_i$ $(i=1,\cdots,n+1)$. The above identity states that $g\phi=\phi A(g)$, $g\,\varepsilon\,K'=SO(n,R)$, that is, the homeomorphism ϕ is equivariant with respect to the isomorphism $A(g)\rightarrow g$ of $A(K')$ onto K'.

Now define $u(x)=(\phi\log)^{-1}x\phi\log$ for $x\,\varepsilon\,K=SO(n+1,R)$. Then u is a homeomorphism (in fact, an isomorphism) of K into aut F. Furthermore, $u(g)=\log^{-1}\cdot\phi^{-1}\cdot g\cdot\phi\cdot\log=\log^{-1}A(g)\log=\lambda(g^{-1})$, $g\,\varepsilon\,K'$. Thus $u(xg)\lambda(g)=u(x)u(g)u(g^{-1})=u(x)$ for all $g\,\varepsilon\,K'$.

It follows now from Theorem 5.1 that $G/C=(K\times F)_K=K/K'\times F$. Thus $Sl(n+1,\mathfrak{R})/Sl(n,\mathfrak{R})$ decomposes, by the covariant fibering of Section 4, into the direct product of an n-sphere and a Euclidean space.

Example 2. By the same reasoning one can prove that

$Sl(n+1,\mathcal{C})/Sl(n,\mathcal{C})\,(SU(n+1)/SU(n))\times$ *Euclidean space.*

These results can be obtained by an elementary geometric argument upon representing elements of our group as a base for $\mathfrak{R}^{n+1}$ (or $\mathcal{C}^{n+1}$ as the case may be).

Example 3. If $G=Sl(3,\mathcal{C})$ and C is the subgroup $Sl(3,R)$, then G/C is the space of a fiber bundle with base $SU(3)/SO(3,\mathfrak{R})$ and a three dimensional Euclidean space F as fiber. The covariant fibering of G/C is not a direct product decomposition. For let us select the Cartan decomposition of $\mathfrak{G}$ (regarded as a real Lie algebra, cf. Section 2) with the maximal subalgebra $\mathfrak{K}$ taken to be the Lie subalgebra of $SU(3)$. Then K, K' are $SU(3)$, $SO(3,\mathfrak{R})$ respectively, and $\lambda(K')$ is equivalent to K' operating on real cartesian space $\mathfrak{R}^3$. The existence of a mapping u of Theorem 5.1 is equivalent to the existence of a continuous mapping u of $SU(3)$ into aut $\mathfrak{R}^3$ such that

$u(xg)g^{-1}=u(x)$, i.e., $u(xg)=u(x)g$ for all x in $SU(3)$, g in $SO(3,\mathfrak{R})$. But such a mapping cannot exist. For if it did, u^{-1}(identity) would be a continuous cross-section of cosets of $SO(3,R)$ in $SU(3)$. This implies that $SU(3)=(SU(3)/SO(3,\mathfrak{R}))\times SO(3,\mathfrak{R})$. But comparing homotopy groups, $\pi_4(SU(3))=0$, whereas $\pi_4(SO(3))\neq 0$ ([13]). Thus $SO(3,R)$ cannot be a direct factor of $SU(3)$ and the given fibering of $Sl(3,\mathfrak{C})/Sl(3,\mathfrak{R})$ is not a product fibering. It should be noted that the isotropy subgroup in this example is connected and *semi-simple*; also the group G is semi-simple.

Example 4. Samelson's example of a Klein space with connected isotropy subgroup but no direct product fibering can be described as follows. Take G to be the subgroup of $Sl(n+1,C)$ whose first column has only zeros below the first row. Let C be the subgroup of G whose coefficients in the second column outside of the $(2,2)$ term consists of zeros. Clearly G and C are connected. G (resp. C) can be identified with the subgroup of projective transformations of the complex projective plane P^n, which keeps a point (resp. two points) invariant. Hence G/C can be identified with the exterior of a point p_0 in P^n. If P^n-p_0 could be factored into a direct product of a compact subset and a Euclidean subspace, we would be able to deform P_∞, the plane at infinity, into a singular cycle which fails to intersect P_∞, contrary to the fact that the singular cycle P_∞ has a non-zero self-intersection. Thus G/C has no direct product fibering. It has, however, a G-covariant fibering. It is noteworthy that G/CR admits a direct product covariant fibering, R being the radical of G. Thus the radical of G plays an essential role in Samelson's example.

Example 5. Let G^* be the connected component of the identity of the Lie group leaving $x_1^2+x_2^2+\cdots+x_n^2-x_{n+1}^2-x_{n+2}^2$ invariant. Let C^* be the analytic subgroup keeping invariant $x_1^2+\cdots+x_n^2-x_{n+1}^2$ together with x_{n+2}. Clearly, G^* and C^* are semi-simple; under the natural imbedding, $SO(n)\times SO(2)$ and $SO(n)$ are maximal compact in G^* and C^* respectively. Let G denote the simply connected covering group of G^*, and let C denote the analytic subgroup corresponding to C^*. Since $G^*=SO(n)\times SO(2)$ $\times$ Euclidean space, the analytic subgroup M in G which covers $SO(n)$ is maximal compact in both G and C. By Theorem 4.1, G/C admits a G-covariant fibering with base space M/M. Hence G/C is homeomorphic to Euclidean space.

Note. The above example was kindly suggested to the author by A. Borel, as an instance of the phenomenon mentioned in the footnotes in Section 6.

6. Aspherical Klein spaces. As an application of Theorem 2.1, we prove

THEOREM 6.1. *An aspherical* (*in dimensions* $0, 1, 2, \cdots$) *Klein space is homeomorphic to Euclidean space.*

Proof. A Klein space is a factor space G/C with G a Lie group and C a closed subgroup. We shall prove the theorem for aspherical factor spaces G/C by induction on the pair of integers $(\dim G, \dim G - \dim C)$ ordered lexicographically. It is convenient to break up the proof into a series of lemmas.

LEMMA 6.1. *Let* $\tilde{G}$ *denote the universal covering group of the connected component* G_0 *of the Lie group* G, *and let* $\tilde{C}$ *be the complete inverse image in* $\tilde{G}$ *of* $C \cap G_0$. *Assume* G/C *is aspherical. Then* $\tilde{G}/\tilde{C} = G/C$. $\tilde{C}$ *is connected and contains a maximal compact subgroup of* G.

Proof. G_0 is a closed normal subgroup of G and G/G_0 is discrete. As a result, G_0C is an open and closed subset of G and G_0C/C is open and closed in G/C. Since G/C is connected, $G_0C/C = G/C$ and therefore

$$\tilde{G}/\tilde{C} = G_0/C \cap G_0 = G_0C/C = G/C.$$

Since $\tilde{G}/\tilde{C}$ is simply connected, $\tilde{C}$ is connected. Let K be any maximal compact subgroup of $\tilde{G}$ which includes a maximal compact subgroup of $\tilde{C}$. By Theorem 3.1, $\tilde{G}/\tilde{C}$ has the same homotopy type as $K/K \cap \tilde{C}$. Since the homotopy groups of all dimensions of $K/K \cap \tilde{C}$ vanish, by the well known theorem of Hurewicz, all the positive dimension homology groups with integral coefficients of $K/K \cap \tilde{C}$ vanish. However, $K/K \cap \tilde{C}$, being a simply connected manifold, is orientable and hence it has a fundamental cycle which is not homologous to zero. It follows immediately that $K/K \cap \tilde{C}$ is a 0-dimensional manifold and being connected, it is a point. Thus $K = K \cap \tilde{C}$. Proof of the lemma is now complete.

LEMMA 6.2. *Let* G *be a Lie group, and let* B, C *be closed subgroups with* $B \supset C$. *If* G/B *and* B/C *are Euclidean* (*spaces topologically*), *then* G/C *is Euclidean.*

Proof. G/C can be considered as a fiber bundle over the case space G/B with fibers B/C, the projection being $p: [xC] \rightarrow [xB]$ and the group of the bundle being the group of left translations of the space of cosets B/C by elements of B. It is known that there always exists a continuous cross-section of a base space into a fiber bundle in case the base space is Euclidean

(Feldbau [5]). In particular, there is a continuous cross-section of G/B into the principal bundle of the bundle G/C. Consequently, the above bundle G/C is the product bundle $G/B \times B/C$. Thus G/C is Euclidean.

LEMMA 6.3. *A maximal proper analytic subgroup (i. e., connected Lie subgroup) of a simply connected Lie group is closed.*

Proof. Let G be the simply connected Lie group and C a maximal proper analytic subgroup. If C is not closed then $\mathcal{C}\ell = G$ and the Lie algebra $\mathfrak{C}$ of C is an ideal in $\mathfrak{G}$ (cf. Section 2; also [7]). Hence C is normal in G and consequently is closed (Pontriagin, *Topological Groups,* p. 279).

LEMMA 6.4. *Let G be a simply connected simple real Lie group, Let C be a closed connected subgroup of G, and let M be a maximal compact subgroup of G. Assume $C \supset M$. Then G/C is homeomorphic to Euclidean space.*

Proof. Let $\mathfrak{G}$, $\mathfrak{C}$, $\mathfrak{M}$ denote the Lie algebras of G, C,.M, respectively. Let $\mathfrak{K} + \mathfrak{E}$ be a Cartan decomposition for the simple Lie algebra $\mathfrak{G}$ (Section 2). Since all maximal compace subgroups of a connected semi-simple Lie group are conjugate and K contains a maximal compact subgroup of G, we may assume that $\mathfrak{M} \supset \mathfrak{K}$ and then $\mathfrak{K} = \mathfrak{M} + \mathfrak{A}$ (direct algebraically) where $\mathfrak{A}$ is abelian and $\mathfrak{M}$ is semi-simple or (0) (cf. [9]). Furthermore, since $\mathfrak{G}$ is simple, ad $\mathfrak{K}$ is irreducible on $\mathfrak{E}$ (cf. [3]).[2] As a result, the center of $\mathfrak{K}$ is at most one-dimensional. In particular, dim $\mathfrak{M}$ is 0 or 1.

We must now distinguish in our argument between the cases $\mathfrak{M} = 0$ and $\mathfrak{M} \neq 0$. If $\mathfrak{M} = 0$, then G is of rank 1, dimension 3, and not compact. In this case, $\mathfrak{G}$ is isomorphic to the Lie algebra of all 2×2 real matrices with trace zero, and we identify $\mathfrak{K}$ and $\mathfrak{E}$ with the skew-symmetric, and symmetric matrices of trace 0 respectively. In this case, it is well known that a one-dimensional subalgebra $\mathfrak{C}$ is conjugate under an inner automorphism to either $\mathfrak{K}$ or a subalgebra of triangular matrices. Hence we can assume if dim $\mathfrak{C} = 1$ that either $\mathfrak{C} = \mathfrak{K}$ or $\mathfrak{C}$ is a subalgebra of triangular matrices. In either case, G/C is homeomorphic to Euclidean space. If dim $\mathfrak{C} = 2$, then as is known, $\mathfrak{C}$ is conjugate to the subalgebra of triangular matrices and $G/C \cap K =$ Euclidean space. (Alternatively, G/C being a one-dimensional simply connected separable space, is Euclidean).

Thus we need only consider the case $\mathfrak{M} \neq 0$. Here, the center of $\mathfrak{K}$ can be described as the centralizer of $\mathfrak{M}$ in $\mathfrak{G}$. For let $\mathfrak{N}$ denote the normalizer $\mathfrak{M}$ in $\mathfrak{G}$, i. e., $N = \{X \mid X \varepsilon \mathfrak{G}, [X, \mathfrak{M}] \subset \mathfrak{M}\}$. Since $\mathfrak{K} \subset \mathfrak{N}$, $\mathfrak{N} = \mathfrak{K} + (\mathfrak{N} \cap \mathfrak{E})$.

---

[2] However, ad $\mathfrak{M}$ need not be irreducible on $\mathfrak{E}$; in fact, there may exist proper subalgebras of $\mathfrak{G}$ other than $\mathfrak{K}$ which include $\mathfrak{M}$, (cf. Example 5 of Section 5).

Since $\mathfrak{N}$ is a subalgebra and $\mathfrak{E}$ is invariant under $\operatorname{ad}\mathfrak{K}$, $\mathfrak{N} \cap \mathfrak{E}$ is invariant under $\operatorname{ad}\mathfrak{K}$. Since $\operatorname{ad}\mathfrak{K}$ is irreducible on $\mathfrak{E}$, $\mathfrak{N} \cap \mathfrak{E} = (0)$ or $\mathfrak{E}$. Thus $\mathfrak{N} = \mathfrak{K}$ or $\mathfrak{N} = \mathfrak{G}$. Since $\mathfrak{G}$ is simple, $\mathfrak{N} = \mathfrak{G}$ implies $\mathfrak{K} = \mathfrak{G} = \mathfrak{N}$. Hence in any case, $\mathfrak{N} = \mathfrak{K}$. Since the centralizer of $\mathfrak{M}$ is in $\mathfrak{N}$, it is in $\mathfrak{K}$ and is central in $\mathfrak{K}$. Conversely, the center of $\mathfrak{K}$ obviously centralizes $\mathfrak{M}$. Thus the center of $\mathfrak{K}$ is the centralizer of $\mathfrak{M}$ in $\mathfrak{G}$, which is of course $\mathfrak{A}$.

We now take up the analysis of the subalgebra $\mathfrak{C}$ under the assumption $\mathfrak{M} \neq (0)$. Inasmuch as $\mathfrak{C} \supset \mathfrak{M}$, $\mathfrak{C} = \mathfrak{M} + (\mathfrak{C} \cap (\mathfrak{A} + \mathfrak{E}))$. Now $\mathfrak{C}$, $\mathfrak{A}$ and $\mathfrak{E}$ being invariant under the completely reducible set of endomorphisms $\operatorname{ad}\mathfrak{M}$, $\mathfrak{C} \cap \mathfrak{E}$ is a subspace of $\mathfrak{C} \cap (\mathfrak{A} + \mathfrak{E})$ invariant under $\operatorname{ad}\mathfrak{M}$. Hence, $\mathfrak{C} \cap (\mathfrak{A} + \mathfrak{E}) = (\mathfrak{C} \cap \mathfrak{E}) + \mathfrak{A}'$, where $\mathfrak{A}'$ is invariant under $\operatorname{ad}\mathfrak{M}$. Since $\dim \mathfrak{A} \leqq 1$, we infer $\dim \mathfrak{A}' \leqq 1$. Since any one-dimensional representation of a semi-simple algebra is the zero representation, $[\mathfrak{M}, \mathfrak{A}'] = 0$. Thus $\mathfrak{A}'$ is in the centralizer of $\mathfrak{M}$ in $\mathfrak{G}$ and hence

$$\mathfrak{A}' \subset \mathfrak{A} \cap \mathfrak{C} \subset \mathfrak{C} \cap \mathfrak{K}, \quad \mathfrak{C} \cap (\mathfrak{A} + \mathfrak{E}) \subset \mathfrak{C} \cap \mathfrak{E} + \mathfrak{C} \cap \mathfrak{K},$$
$$\mathfrak{C} \subset \mathfrak{M} + \mathfrak{C} \cap (\mathfrak{A} + \mathfrak{E}) \subset \mathfrak{C} \cap \mathfrak{K} + \mathfrak{C} \cap \mathfrak{E}.$$

We have proved therefore that $\mathfrak{C} = \mathfrak{C} \cap \mathfrak{K} + \mathfrak{C} \cap \mathfrak{E}$. Applying Theorem 2.1 we have that $G/C = (M \times W)_{M'}$ where $M' = C \cap M = M$, and W is Euclidean. Thus G/C is homeomorphic to the space of a fiber bundle with Euclidean fiber W and with a base space $M/M' = M/M$, which is a point. Therefore, G/C is homeomorphic to Euclidean space.

We now proceed to prove Theorem 6.1. By Lemma 6.1, we may assume that G is simply connected and that C is an analytic subgroup containing a maximal compact subgroup of G. We prove by induction on the pair of integers $(\dim G, \dim G/C)$ ordered lexicographically, that: *if G is a simply connected analytic group and C is a closed (connected) analytic subgroup which contains a maximal compact subgroup of G, then G/C is Euclidean space, topologically.*[3] This induction hypothesis is certainly true for the pair $(0, 0)$. Assume it true for pairs less than $(\dim G, \dim G/C) = (r, n)$. Let G_1 be a closed proper maximal analytic subgroup of G which contains C. Inasmuch as an analytic group is topologically the direct product of a maximal compact subgroup and a Euclidean space and furthermore, all maximal compact subgroups of an analytic group are conjugate, any maximal compact subgroup of simply connected Lie group is simply connected. G_1 is thus the direct product of a simply connected compact subgroup and a

---

[3] This result is claimed for semi-simple groups by L. Calabi in Rend. di Mat. di Univ. di Roma, Ser. 5, vol. 7 (1952), p. 326, Cor. 2. However, his proof has a central gap and would not apply to the case in Example 5 of Section 5.

Euclidean space and is hence simply connected. We can therefore apply the induction hypothesis to G/G and G/C if G contains C properly. By Lemma 6.2, G/C is Euclidean in this case. We need only consider the case $G_1 = C$, i.e., C is a maximal closed analytic subgroup of G. By Lemma 6.3, C is a maximal proper analytic subgroup of G.

Let now $\mathfrak{N}$ be any maximal ideal of $\mathfrak{G}$ which is not $\mathfrak{G}$. CN is then an analytic subgroup containing C. Thus $G = CN$ or $CN = C$. If $G = CN$, then $G/C = N/N \cap C$. By Lemma 6.1, we may assume that N and $N \cap C$ fulfill the hypotheses of our induction hypotheses (as a matter of fact, N *is* simply connected and $N \cap C$ *is* connected and contains a maximal compact subgroup of N). By our induction hypothesis $N/N \cap C$ is Euclidean, and thus G/C is Euclidean if $CN = G$.

It remains only to consider the case $CN = C$, i.e., $N \subset C$. Here $G/C = G/N/C/N$ and if $N \neq (0)$, the induction hypothesis yields the result that G/C is Euclidean. We are reduced to considering the case $\mathfrak{N} = 0$, i.e., the only ideals of $\mathfrak{G}$ are $\mathfrak{G}$ and (0). Thus G is either isomorphic to the additive group real numbers or G is a (non-abelian) simple Lie group. But by Lemma 6.4, G/C is Euclidean in this case. Our induction hypothesis has therefore been verified for the pair (r, n) and Theorem 6.1 is now proved.

7. Uniqueness of covariant fiberings. In Section 4 we defined "covariant fibering of a Klein space" and in Section 5 we saw by an example that covariant fiberings are not always trivial (i.e., direct product decompositions) even when the isotropy subgroup is connected and semi-simple. In order to understand better the significance of such an assertion, we must inquire into the uniqueness of covariant fiberings. Our concern is with the *decomposition* of the homogeneous space into fibers rather than the structure group of the fiber bundle. To that end we regard the fibered homogeneous spaces as fiber bundles whose structural group is the topological group of all homeomorphisms of the Euclidean fiber onto itself (cf. [0]).

In seeking a uniqueness theorem for covariant fiberings of a homogeneous H, we avoid the knotty problem of cofactorizations of a topological space by comparing only G-covariant fiberings. That is, we are concerned here only with fiberings that are suitably related to the operations of G on H.

Throughout this discussion, G will denote a fixed Lie group which operates transitively on a homogeneous space H. C_p will denote the isotropy subgroup of the point p in H, i.e., the subgroup of G which keeps p fixed. We assume that C_p is connected. We will consider only G-covariant fiberings of H.

Let $\mathfrak{F}$ be a fibering of H.

Definition. A *transversal* M to $\mathfrak{F}$ is a maximal compact subgroup of G which permutes the fibers of $\mathfrak{F}$. A *base point* of the transversal M is a point p whose isotropy subgroup C_p intersects M in a maximal compact subgroup of C_p. The *F-isotropy subgroup* of M is the subgroup M_F which keeps the fiber F invariant. If p is a base point of M, we call $M \cap C_p = M_p$ the *M-isotropy* subgroup of p.

We recall that a G-covariant fibering $\mathfrak{F}$ of H is a decomposition of the differential structure of H into Euclidean fibers for which there exists a transversal M with the properties:

1) The F-isotropy subgroup M_F is equivalent to a linear group on F for some fiber F.

2) M permutes the fiber transitively.

As a consequence of 2), M_F is equivalent to a linear group on F for *every* fiber F.

PROPOSITION 7.1. *Let M be a transversal of the covariant fibering $\mathfrak{F}$, and let M_F be the F-isotropy subgroup of M. Then M_F is equivalent to a linear group on the Euclidean fiber F and M permutes the fibers of $\mathfrak{F}$ transitively.*

Proof. Let G_1 be the connected component of the identity in the subgroup of G which permutes the fibers of $\mathfrak{F}$. Let G_F be the connected component of the identity in the subgroup of G_1 which keeps the fiber F invariant. Both G_1 and G_F are closed subgroups of G.

M_F is a maximal compact subgroup of G_F. For let N_F be a maximal compact subgroup of G_F which contains M_F and let N be a maximal compact subgroup of G_{11} which contains N_F. Since G_1 and G_F are connected Lie groups, there is an inner automorphism $\Gamma_g : x \to g \times g^{-1}$ with g in G, which carries M onto N and M_F into N_F (by Theorem 2.3). Since the transformation $h \to g \cdot h$ (g in G, h in H) of H onto itself is a fiber-preserving homeomorphism, Γ_g takes M_F, the F-isotropy subgroup of M, onto N_{gF}, the gF-isotropic subgroup of N. In turn, N_{gF} is conjugate under an inner automorphism of N to N_F. Thus M_F is conjugate to N_F under an inner automorphism of G_1. Since $M_F \subset N_F$, we have $\dim M_F = \dim N_F$, M_F is open and closed in N_F, and hence $M_F = N_F$. Thus, M_F is a maximal compact subgroup of G_F.

Now F being a covariant fibering, there is a transversal M' to F whose F-isotropic subgroup M'_F operate on F like a linear group and which permutes the fibers transitively. Since, by the foregoing, M'_F is a maximal compact subgroup of the connected Lie group G_F, it is conjugate to M_F under an inner automorphism Γ_g with g in G_F. Since g keeps F invariant, M_F is equivalent to a linear group on $gf = F$. Moreover, M' being a maximal compact subgroup of the connected Lie group G_1, is conjugate to M by an inner automorphism of G_1. Hence M also permits the fibers transitively.

We now turn to a proof of the uniqueness of G-covariant fiberings of H up to equivalence in the sense that there exists a one-to-one bundle map between any two.

A covariant fibering $\mathfrak{F}$ is completely determined in view of Proposition 7.1 by the specification of a transversal M and a single fiber F. We shall employ the notation (M,F) for the covariant fibering $\mathfrak{F}$ which has M as transversal and F as fiber.

LEMMA 7.1. *Let (M,F) be a G-covariant fibering. Then (gMg^{-1}, gF) is a G-covariant fibering for any g in G and the map $x \to gx$ of H is a one-to-one bundle map of (M,F) onto (gMg^{-1}, gF).*

The verification of Lemma 7.1 is trivial.

LEMMA 7.2. *Let M be a transversal to the covariant fibering $\mathfrak{F}$, let p be a base point of M and let F be the fiber through p. Then $M_p = M_F$,* i.e., *$M \cap C_p$ is the subgroup of M which keeps F invariant.*

Proof. If we make use of the fact that M_F is equivalent to a linear group on F, then we infer that it keeps a point q fixed, i.e., $M_q = M_F$. Inasmuch as M_q is a maximal compact subgroup of C_q and C_q is conjugate to C_p under an inner automorphism, we find that M_p and M_q are conjugate under an inner automorphism since each is a maximal compact subgroup of conjugate connected Lie subgroups. M_p is a subgroup of the connected group M_F and, having the same dimension as M_F, must coincide with M_F.

LEMMA 7.3. *Let p be a base point of the transversal M. Then the orbit $M(p)$ has no tangent vector which is also tangent to a fiber.*

Proof. Let F_x denote the fiber through the point x, and let V_x denote the tangent space to F_x at x. Let U_x denote the tangent space to the orbit $M(p)$ at the point x in $M(p)$. It is clearly sufficient to prove that $U_p \cap V_p = (0)$. By the well known first fundamental theorem of Sophus Lie, the connected group M, considered as a transformation group operating (effectively) on H is generated by a linear family of infinitesimal transformations $\mathfrak{M}$ which can be identified with the Lie algebra of M. Let $\mathfrak{M}'$ denote the subalgebra of $\mathfrak{M}$ which generates the subgroup M_p. We denote by $\mathfrak{M}_x$

and $\mathfrak{M}'_x$ the element of contact which the distributions $\mathfrak{M}$ and $\mathfrak{M}'$ assign to the point x. Clearly, $\mathfrak{M}_p = U_p$ and $\mathfrak{M}'_p = (0)$.

Suppose now that $\mathfrak{M}_p \cap V_p \neq (0)$. Then there is an infinitesimal transformation X in $\mathfrak{M}$, but not in $\mathfrak{M}'$, such that X_p is in V_p. Now by definition of the exponential of an infinitesimal transformation, $X_{\exp X(q)}$ is the tangent vector to the parametrized path $\exp tX(p)$ at the point $t = 1$; i. e., for any function f analytic on H around $\exp X(q)$, $(Xf)_{\exp X(q)} = df(\exp tX(q))/dt$ at $t = 1$. Hence, X is tangent to the path $p(t)$ at all values of t, where $p(t)$ denotes the path $\exp tX(p)$.

On the other hand, $X_{p(t)}$ is the image of X_p under the differential of the transformation $\exp tX$, in view of the fact that $\exp tX$ carries the path $\exp sX$ (s varying) into the path $\exp(t+s)X(p)$. Since $\exp tX$ is a transformation in M, it permutes the fibers and hence its differential permutes the V_x with x in the orbit M. As a result, $X_{p(t)}$ is in $V_{p(t)}$ for all t and the path $\exp tX(p)$ is tangent to a fiber for each t.

Since a covariant fibering decomposes the differential structure of H, we may introduce a coordinate system $y_1, \cdots, y_n$ compatible with the differentiable structure of H such that in a neighborhood of p the fibers are slices $y_1 = \text{constant}, \cdots, y_f = \text{constant}$.

Since $p(t)$ is tangent to a fiber for each t, we have $dy_i(p(t))/dt = 0$ $(i = f, \cdots, n)$, and hence $y_i(p(t)) = \text{constant} = y_i(p(0))$. It follows immediately that $p(t)$ lies in the slice F_p. Since $\exp tX$ permutes fibers, it must be in M_F. By Lemma 7.2, $\exp tX$ is in M_p for each t and hence $X_p = 0$, contrary to hypothesis. Consequently, $U_p \cap V_p = (0)$.

LEMMA 7.4. *Let N be a compact group of transformations on a Euclidean space F which is equivalent to a linear group. Let p be any fixed point of N, let V be the tangent space to F at p. For each transformation g in N, let dg denote the differential of g at p. Then N is equivalent to the linear group dN acting on V with p corresponding to the zero vector and any element g of N corresponding to dg.*

Proof. By hypothesis, there is a coordinate system $x_1, \cdots, x_n$ on F (not necessarily differentiable) with respect to which M is a group of linear transformations. Since p is fixed under M, upon replacing each i-th coordinate function by $x_i - x_i(p)$, we arrive at a coordinate system in which M is a linear group and p is the origin. Thus without loss of generality we may assume that p is the origin of the coordinate system $x_1, \cdots, x_n$. By a result of Bochner (cf. [1]) there is a *neighborhood* $\mathcal{O}$ of the point p and differentiable mapping ϕ_1 of $\mathcal{O}$ onto a neighborhood of zero in V such that ϕ_1 is equivariant with respect to the isomorphism d (that d is an isomorphism is well known; for expressing elements of the kernel N_d of d as linear parts plus terms of higher order, it is seen that no iterate of a transformation in N_d is of finite order and hence N_d, containing no non-unit elements of finite order, reduces to the unit).

We regard F as a linear space with $x_1, \cdots, x_n$ as a base of linear functions. Let $X_1, \cdots, X_n$ be a set of base vectors of F which lie in the neighborhood $\mathcal{O}$. By Bochner's result, $dg\phi_1(X_i) = \phi_1 g(X_i)$ $(i = 1, \cdots, n)$. We define the linear mapping ϕ of F onto V by the formula $\phi(\Sigma c_i X_i) = \Sigma c_i \phi(X_i)$. By linearity we get $dg\phi(X) = \phi g(X)$, all X in F_1. Hence, ϕ is equivariant with respect to d and N is equivalent to dN on V.

Note. ϕ may be non-differentiable with respect to the given differential structures of F.

LEMMA 7.5. *Let M be a transversal of a G-covariant fibering, F a fiber, p a point in F that is fixed under M_F, and C_p the isotropy subgroup of p. Then M_F operating on F is equivalent to* $\mathrm{Ad}\, M_F$ *on $\mathfrak{G}/\mathfrak{M} + \mathfrak{C}_p$ with p corresponding to zero, and m corresponding to the $\mathfrak{G}/\mathfrak{M} + \mathfrak{C}_p$ part of* $\mathrm{Ad}\, m$ *for m in M_p.*

Proof. Let H_p be the tangent space to H at p, let U_p be the tangent space to the orbit M_p at p and let V_p be the tangent space at p to the fiber F through p. Then H_p and U_p are the images of the Lie algebras $\mathfrak{G}$ and $\mathfrak{M}$ under the differential of the mapping $\pi : g \to g \cdot p$ of G onto the homogeneous space H. Now for any c in C_p and x in G, $\pi(cx) = \pi(cxc^{-1})$. Inasmuch as $c\pi(x) = \pi(cx)$ by definition of the operation of G on H, we have $c \cdot \pi = \pi \cdot \Gamma_c$, c in C_p. On taking the differential of this identity at the identity, we get $dc_p \cdot d\pi = d\pi \cdot \mathrm{Ad}\, c$, since $\mathrm{Ad}\, c$ is $d\Gamma_c$ at the identity. Denoting by dC_p the differentials of the transformations of H which are in C_p, we see that dC_p is equivalent to the operation of $\mathrm{Ad}\, C_p$ on $H_p = d\pi(G)$ which can be identified with $\mathfrak{G}/\mathfrak{C}_p$. In particular, the subgroup dM_p of differentials from M_p is equivalent to $\mathrm{Ad}\, M_p$ operating on $\mathfrak{G}/\mathfrak{C}_p$. By Lemma 7.3, V_p is complementary to U_p in H_p and hence dM_0, operating on V_p is equivalent to dM_p operating on H_p/U_p. Inasmuch as U_p corresponds to $\mathfrak{M} + \mathfrak{C}_p/\mathfrak{C}_p$ under the given identification of H_p with $\mathfrak{G}/\mathfrak{C}_p$, we conclude that dM_p ouerates on V_p as $\mathrm{Ad}\, M_p$ operates on $\mathfrak{G}/\mathfrak{C}_p/\mathfrak{M} + \mathfrak{C}_p/\mathfrak{C}_p = \mathfrak{G}/\mathfrak{M} + \mathfrak{C}_p$. Applying Lemma 7.4 and noting that $M_p = M_F$, by Lemma 7.2, we obtain the desired conclusion.

THEOREM 7. *Let G be a connected Lie group which operates transitively on the homogeneous space H with connected isotropy. Then any two G-*

covariant fiberings of H are equivalent decompositions, that is, there is a one-to-one bundle map of one onto the other.

Proof. By Lemma 7.1 we may restrict our attention to G-covariant fiberings with the same transversal M. Let therefore (M,F_1) and (M,F_2) be the two covariant fiberings. Inasmuch as the property of being a base point to M depends only on M, there is no generality lost in assuming that F_1 and F_2 contain the same base point p. By Lemma 7.2, $M_{F_1}=M_p=M_{F_2}$. By Lemma 7.5, there is a homeomorphism ϕ of F_1 onto F_2 such that $\phi(mf)=m\phi(f)$ for m in M_p and f in F_1. It follows immediately that the mapping $mf\rightarrow m\phi(f)$ for m in M, f in F_1, is a well defined homeomorphism which is a bundle map of (M,F_1) onto (M,F_2).

By an *isomorphism* of a fiber bundle we mean a one-to-one (bicontinuous) bundle map of it. Let M_i be a group of transformations on a bundle B_i $(i=1,2)$, and let ϕ be an isomorphism of B_1 onto B_2. We say ϕ send M_1 onto M_2 if $\phi M_1\phi^{-1}=M_2$.

LEMMA 7.6. *Let M be a transversal to the G-covariant fiberings $\mathfrak{F}$. The base points of M are permuted transitively by the automorphism of $\mathfrak{F}$ which keep M invariant.*

Proof. Let p, p be any two base points to M and let F_1, F_2 be the fibers through them. Since (Prop. 7.1) M permutes the fibers transitively, we may assume without loss of generality that $F_1=F_2=F$. Hence, $M_{F_1}=M_{F_2}=M_F$.

If a linear space V and p_1, p_2 are fixed points of L, then the mapping $v\rightarrow v+(p_1-p_2)$ of V is equivariant with respect to the identity automorphism. Since M_F is equivalent to a linear group on F, there is a homeomorphism ϕ of F onto itself such that $\phi(mf)=m\phi(f)$ for all m in M_F, f in F. As a result, $mf\rightarrow m\phi(f)$ with m in M and f in F is an automorphism of $\mathfrak{F}$ which carries p_1 into p_2. Since this automorphism keeps M invariant, the desired conclusion holds.

The above uniqueness result can be strengthened.

PROPOSITION 7.2. *Let F_i be a G-covariant fibering of H with transversal M_i and base point p_i $(i=1,2)$. Then there is an isomorphism of F_1 onto F_2 which carries p_1 onto p_2 and the transversal M_1 onto M_2.*

Proof. By Lemma 7.1, we reduce to the case that $M_1=M_2$. Let F_1, F_2 denote the fibers of $\mathfrak{F}_1$, $\mathfrak{F}_2$ respectively that pass through p_1. From Lemma 7.5 it follows (just as in the proof of Theorem 7) that there is a mapping θ of F_1 onto F_2 such that the map $\phi_1\colon mf\rightarrow m\theta(f)$ (m in M, f in F_1) is an isomorphism of $\mathfrak{F}_1$ onto $\mathfrak{F}_2$. Clearly, $\phi^{-1}m\phi=m$ for all m in M.

On the other hand, by Lemma 7.6, there is an automorphism ϕ_2 of $\mathfrak{F}_2$ which carries p_1 into p_2 and keeps M invariant. Hence $\phi_2\phi_1$ is an isomorphism of $\mathfrak{F}_1$ onto $\mathfrak{F}_2$ which sends p_1 onto p_2 and keeps M invariant. Proof of the proposition is now complete.

Remarks. In our uniqueness discussion for covariant fiberings, we made repeated use of the hypothesis that M_F is equivalent to a linear group on the fiber F. There is much ground for the suspicion that this hypothesis is superfluous in the sense that it is always true. That is, there is the well known conjecture of Samelson that any compact group of differentiable automorphisms of Euclidean space is equivalent to a linear group. In our case, even a weaker theorem is needed to conclude that M_F is equivalent to a linear group on F. Namely, let K be a compact subgroup of a Lie group G, and let F be a closed Euclidean (differentiable) subspace of a space H on which G acts transitively. If K keeps F invariant, it is equivalent to a linear group on F.

This conjecture appears to be difficult to prove even in the case that G is the group of rigid motions in Euclidean space and K is a subgroup of the orthogonal group.

THE JOHNS HOPKINS UNIVERSITY
AND
INSTITUTO DE MATEMATICA PURAE APLICADA, RIO DE JANEIRO.

REFERENCES.

[0] R. Arens, "Topologies for homeomorphism groups," *American Journal of Mathematics*, vol. 68 (1946), pp. 593-610.

[1] S. Bochner, "Compact groups of differentiable transformations," *Annals of Mathematics*, vol. 46 (1945), pp. 377-381.

[2] A. Borel and J. P. Serre, "Impossibilité de fibrer un espace Euclidien par des fibrés compactes," *Comptes Rendus de l'Académie des Sciences (Paris)*, vol. 230 (1950), pp. 2258-2260.

[3] E. Cartan, "Groupes simples clos et ouverts et géométrie Riemannienne," *Journal de Mathématiques Pures et Appliquées*, vol. 8 (1929), pp. 1-33.

[4] C. Chevalley, "On the topology of solvable groups," *Annals of Mathematics*, vol. 42 (1941), pp. 668-675.

[5] Feldbau, "Sur la classification des espaces fibrés," *Comptes Rendus de l'Académie des Sciences (Paris)*, vol. 208 (1939), pp. 1621-1623.

[6] K. Iwassawa, "On some types of topological groups," *Annals of Mathematics*, vol. 50 (1949), pp. 507-558.

[7] A. Malcev, "On the theory of Lie groups in the large," *Recueil Mathématique (Matematicheskii Sbornik)*, vol. 16 (58) (1945), pp. 163-189.

[8] D. Montgomery, "Simply connected homogeneous spaces," *Proceedings of the American Mathematical Society*, vol. 1 (1950), pp. 467-469.

[9] G. D. Mostow, "A new proof of E. Cartan's theorem on the topology of semi-simple groups," *Bulletin of the American Mathematical Society*, vol. 55 (1949), pp. 969-980.

[10] ———, "Some new decomposition theorems for semi-simple groups," *Memoirs of the American Mathematical Society*, (to appear).

[11] ———, "Factor spaces of solvable groups," *Annals of Mathematics*, vol. 60 (1954), pp. 1-27.

[12] A. Shapiro, "Cohomologie dans les espaces fibrés," *Comptes Rendus de l'Académie des Sciences (Paris)*, vol. 231 (1950), pp. 206-207.

[13] N. Steenrod, *The Topology of Fiber Bundles*, Princeton University Press, 1951.

[14] J. H. C. Whitehead, "On the decomposition of an infinitesimal group," *Proceedings of the Cambridge Philosophical Society*, vol. 32 (1936), pp. 229-237.

ON THE LIE AND JORDAN RINGS OF A SIMPLE ASSOCIATIVE RING.*

By I. N. HERSTEIN.

Given any associative ring A we can form, using its operations and its elements, two new rings. These use the elements of A and the addition as defined in A, but new multiplications are introduced to render them rings, albeit not necessarily associative rings. The first of these, the *Lie ring* A^L of A uses a multiplication defined by $[a, b] = ab - ba$ for any $a, b \varepsilon A$ where ab is the ordinary associative product of elements in A. The second of these, the *Jordan ring* of A, A^J, has its multiplication defined by $a \circ b = ab + ba$ for any pair of elements a, b in A.

Being defined in a manner so decidedly dependent on the associative product of A, it is natural to expect that an intimate relationship should exist between the structure of these two new rings and that of A. In this paper we study one phase of this relationship, namely the connection between the ideal structure of A as an associative ring with the ideal structure of A^L and A^J as Lie and Jordan rings respectively. To be more specific, we investigate how simplicity of A as an associative ring reflects into analogous properties of A^L and A^J.

When we say that U is an ideal of A^J, or, equivalently, when we say that U is a Jordan ideal of A, we mean that U is an additive subgroup of A and that for any $x \varepsilon U$ and any $y \varepsilon A$, $x \circ y = xy + yx$ is an element of U. We similarly define Lie ideals of A and ideals of A^L.

Although the main results of this paper deal with the case in which A is a simple ring, many of the other results do not require the assumption of simplicity in order to remain valid; so, unless otherwise stated, we make no assumption of simplicity for A.

1. The Jordan ring of A.

We begin with

LEMMA 1. *If U is a Jordan ideal of A, and if a and b are elements of U, then for any x in A $(ab + ba)x - x(ab + ba)$ is an element of U.*

Proof. Since a is in U, a Jordan ideal of A, then for all $x \varepsilon A$, $a(xb - bx) + (xb - bx)a$ is also in U. But

---

* Received August 31, 1954.

$$a(xb-bx)+(xb-bx)a$$
$$=\{(ax-xa)b+b(ax-xa)\}+\{x(ab+ba)-(ab+ba)x\}$$

and since $b \varepsilon U$, the first term on the right hand side of this equation is in U; this, coupled with $a(xb-bx)+(xb-bx)a$ in U leads to $x(ab+ba)-(ab+ba)x \varepsilon U$, which is the lemma we set out to prove.

LEMMA 2. *If A is a simple ring of characteristic different from 2, and if U is a Jordan ideal of A, $U \neq A$, then any element a of U with the property that $ax-xa \varepsilon U$ for all x in A must necessarily satisfy $a=0$.*

Proof. By assumption, $ax-xa \varepsilon U$ for all $x \varepsilon A$. On the other hand, since U is a Jordan ideal of A, $ax+xa \varepsilon U$ for all x in A. Adding, we obtain that $2ax \varepsilon U$ for all $x \varepsilon A$. Since A is of characteristic $\neq 2$, $2A=A$, so we have that $ax \varepsilon U$ for all $x \varepsilon A$. However, for all y, since $ax \varepsilon U$, a Jordan ideal of A, $(ax)y+y(ax) \varepsilon U$; and since by the above $axy \varepsilon U$, we must have that $yax \varepsilon U$. In this way $AaA \subset U$. Since A is simple, $AaA=A$ unless $a=0$. Thus the lemma is established.

THEOREM 1. *If A is a simple ring of characteristic different from 2, then A^J is a simple Jordan ring.*

Proof. Suppose that $U \neq A$ is a Jordan ideal of A. For any $a, b \varepsilon U$ and any $x \varepsilon A$, by Lemma 1, $(ab+ba)x-x(ab+ba)$ is in U. Since $ab+ba$ is an element of U, by Lemma 2, $ab+ba=0$. In particular, $2a^2=0$, and so $a^2=0$ for any a in U. If $a \varepsilon U$, then for all $x \varepsilon A$, $ax+xa \varepsilon U$; so by the above, with $b=ax+xa$, we have that $a(ax+xa)+(ax+xa)a=0$; since $a^2=0$ we are left with $2axa=0$, and so, $axa=0$, for all $x \varepsilon A$. But then aA is a nilpotent right ideal of A, which is, of course, impossible in a simple ring unless $a=0$. Thus $U=(0)$ and we have proved that A^J is a simple Jordan ring.

(Note: the argument used in the proof of the above theorem actually can be refined to prove somewhat more, namely: if A is an associative ring such that no homomorphic image of A possesses nilpotent right ideals, then any Jordan ideal of A is at the same time an ordinary ideal of A. I owe this remark to Professor Irving Kaplansky.)

2. The Lie ring of *A*. For $a \varepsilon A$, by $[a, A]$ we shall mean the set $\{ax-xa \mid x \varepsilon A\}$. $[A, A]$ shall then represent the subgroup of A generated by all the commutators $xy-yx$ for all x, y in A. When we refer to an ideal of A we shall mean an ordinary, two-sided ideal of A under the associative multiplication of A.

A ring R, all of whose elements are nilpotent, is said to be *locally nilpotent* if the subring generated by any finite set of elements of R is nilpotent. An ideal of R is said to be locally nilpotent if, as a ring, it is locally nilpotent.

We proceed to

THEOREM 2. *Let A be a ring with no non-zero locally nilpotent ideals. Suppose that in A, $2x = 0$ implies that $x = 0$. Suppose further that U, an associative subring of A, is also a Lie ideal of A. Then either U contains a non-zero ideal of A or U is contained in the center of A.*

Proof. Suppose first that U, as a subring of A, is not commutative. Then for some x, y in U, $xy - yx \neq 0$. Since U is a Lie ideal of A and since $x \in U$, $x(ys) - (ys)x$ is in U for every s in A. But $x(ys) - (ys)x = (xy - yx)s + y(xs - sx)$. Since $x \in U$ and since U is a Lie ideal of A, $xs - sx$ is in U; since U is, in addition, a subring of A and since y is also in U, $y(xs - sx)$ is in U. Thus from the above equation we have that $(xy - yx)s \in U$ for all $s \in A$. Thus for any $r \in A$, since U is a Lie ideal of A,

$$r(xy - yx)s - (xy - yx)sr = r\{(xy - yx)s\} - \{(xy - yx)s\}r \in U.$$

Knowing that $(xy - yx)sr \in U$ for all $r, s \in A$ from the argument above, we can now conclude that $r(xy - yx)s$ is in U for every r and s in A. In this way the ideal $A(xy - yx)A \subset U$. Hence if the theorem were false, $A(xy - yx)A = (0)$. This would force $(xy - yx)A$ to be a nilpotent ideal of A, which by assumption would mean that $(xy - yx)A = (0)$. $xy - yx$ would then become an absolute zero divisor, which is again impossible in a ring free from locally nilpotent ideals. So we have been forced into a contradiction. We must therefore suppose that U is a commutative subring of A.

We now propose to show that if U is a commutative ring then it must lie in the center of A. We assume that there is an $x \in U$ and an $s \in A$ so that $y = xs - sx \neq 0$. Since $xs^2 - s^2x$ is in U, a commutative ring, $x(xs^2 - s^2x) = (xs^2 - s^2x)x$. However, from $xs^2 - s^2x = (xs - sx)s + s(xs - sx)$ we have that $x\{(xs - sx)s + s(xs - sx)\} = \{(xs - sx)s + s(xs - sx)\}x$. Since $xs - sx$ is also in U, it must commute with x; making use of this in the last equation above we obtain that $2(xs - sx)^2 = 0$. By assumption on A this leads to $(xs - sx)^2 = 0$. $y = xs - sx \neq 0 \in U$, so for all $t \in A$, $(yt - ty)^2 = 0$. Since $y^2 = 0$, left multiplying $(yt - ty)^2 = 0$ by y and right multiplying it by t, we obtain that $(yt)^3 = (0)$ for all $t \in A$. By a result of Levitzki [1] yA is a locally nilpotent right ideal. But then, by another result of Levitzki [2] AyA is a locally nilpotent ideal of A. In this way we arrive at a contradiction. So $y = 0$, and x is in the center of A. That is, U is contained in the center of A.

It is generally well known, although to this author's knowledge nowhere in the literature, that a simple ring can not be locally nilpotent. For the sake of completeness we record a proof of this here, since we must make use of it in the next theorem. So suppose that A is simple and locally nilpotent. Thus if $a \neq 0 \varepsilon A$, $AaA = A$. Thus for some $r_1, \cdots r_n, s_1, \cdots s_n$ in A

$$a = r_1 a s_1 + r_2 a s_2 + \cdots + r_n a s_n.$$

The subring T generated by $r_1, r_2, \cdots, r_n, s_1, \cdots, s_n$ is nilpotent, say $T^k = 0$. If we substitute the expression above for a in the right hand side for each a occurring we obtain $a = r_1' a s_1' + \cdots + r_t' a s_t'$, where the r_i' and s_i' are in T^2. Continuing this way we obtain such an expression for a where the multipliers on the right hand side come from T^{2^k}, and so are all 0; but then $a = 0$, a contradiction. This proves the assertion that a simple ring can not be locally nilpotent.

With this now established we have as an immediate corollary to Theorem 2

THEOREM 3. *If A is a simple ring of characteristic different from 2, then any proper Lie ideal of A which is at the same time a subring of A must lie in the center of A.*

LEMMA 3. *Let A be any associative ring, and let U be a Lie ideal of A. Let $T(U) = \{t \varepsilon A \mid [t, A] \subset U\}$. Then $T(U)$ is both a Lie ideal and a subring of A. Moreover $U \subset T(U)$.*

Proof. $T(U)$ is trivially a Lie ideal of A. Suppose that $a, b \varepsilon T(U)$, $c \varepsilon A$. Then $(ab)c - c(ab) = \{a(bc) - (bc)a\} + \{b(ca) - (ca)b\}$, and since both a and b are in $T(U)$, the right hand side of this equation is in U. Thus $[ab, A] \subset \mathrm{U}$, whence $ab \varepsilon T(U)$ and $T(U)$ is a subring of A. Since U is a Lie ideal of A, $U \subset T(U)$.

We are now in position to prove

THEOREM 4. *Let A be a simple ring of characteristic different from 2. Let U be a Lie ideal of A. Then either U is contained in the center of A or $U \supset [A, A]$.*

Proof. $T(U)$ is both a subring and a Lie ideal of A. By Theorem 3 either $T(U)$ is contained in the center of A or $T(U) = A$. If $T(U)$ is contained in the center of A, then so is U since $U \subset T(U)$. If, on the other hand $T(U) = A$, then by the very definition of $T(U)$, $[A, A] \subset U$.

3. Case of characteristic 2. Suppose that A is a simple ring of characteristic 2. Theorem 4 need no longer be true in this case. In fact, if A is the set of 2×2 matrices over a field Z of characteristic 2, or if A is a division algebra of dimension 4 over its center, a field of characteristic 2, then the theorem is definitely false. We propose to show that these constitute the only exceptions to our main result.

In the chain leading from Theorem 3 through Theorem 4, the characteristic of A does not enter. Thus we will consider those situations in which Theorem 3 fails to hold true in case A is of characteristic 2.

We suppose that A is a simple ring of characteristic 2. Suppose that U is both a Lie ideal and a subring of A and that $U \neq A$. From the Jacobi identity we obtain, as we did in the proof of Theorem 2, that if $x, y \varepsilon U$ and $s \varepsilon A$, then $(xy + yx)s \varepsilon U$. That is, $(xy + yx)A \subset U$. Hence as before, $A(xy + yx)A \subset U$. By the simplicity of A, $A(xy + yx)A$ is all of A unless $xy + yx = 0$; we are thus forced to assume that $xy + yx = 0$ for all x and y in U. If $U \subset Z$, the center of A, we have nothing left to prove. Suppose now that there exists an a in U and a is not in Z. Thus $as + sa \varepsilon U$ for all $s \varepsilon A$. Consequently $a(as + sa) + (as + sa)a = 0$ by the above; that is, $a^2 s = sa^2$ for all $s \varepsilon A$; so $a^2 \varepsilon Z$. Since $ar + ra \varepsilon U$, we also have that $(ar + ra)^2 \varepsilon Z$ for all $r \varepsilon A$. Should $Z = (0)$, then by the argument used in Theorem 2, Levitzki's theorems would apply, and would lead to a contradiction with the simplicity of A. So we can assume that $Z \neq (0)$. We can likewise assume that $a^2 \neq 0$ (for if $(ax + xa)^2 = 0$ for all x and if $a^2 = 0$, then Levitzki's theorems could again be used) for some $a \varepsilon U$, $a \notin Z$.

$Z \neq (0)$ is a field of characteristic 2. Since Z is not trivial, A is a primitive ring, and so is a dense ring of transformations on a vector space V over a division ring D. We want to first show that V is at most 2-dimensional over D, and from this to deduce that A is 4-dimensional over Z. If we extend Z to a field Z' we obtain a new simple ring A' and the dimension of A' over Z' is the same as the dimension of A over Z. Also, since $ax + xa$ and $ay + ya$ are in U for x, y in A, these two commutators always commute; from this it follows $(aw + wa)^2 \varepsilon Z'$ for all $w \varepsilon A'$. So all the properties of A carry over to A'. We make a special choice for Z', namely, since $a^2 = \lambda \varepsilon Z$, we define $Z' = Z(\sqrt{\lambda})$. Thus in A', $a^2 = \mu^2$, $\mu \varepsilon Z'$; the element $a' = a/\mu$ then satisfies $(a'w + wa')^2 \varepsilon Z'$ for all $w \varepsilon A'$ and $(a')^2 = 1$. All this discussion has achieved is that without loss of generality we may assume that in A there exists an element a, $a \notin Z$, $a^2 = 1$, $(ax + xa)^2 \varepsilon Z$ for all $x \varepsilon A$. On this basis we shall show that A is 4-dimensional over Z. Since $(a + 1)^2 = 0$, $a + 1 \neq 0$, A has zero divisors, so is not a division ring. As we noted earlier, A is a dense ring of linear transformations on a vector space V over a division ring D.

Let us first suppose that we can find v_1 and v_3 in V so that $v_1, v_2 = v_1 a$, $v_3, v_4 = v_3 a$ are linearly independent over D. Since A is dense on V there exists an $x \varepsilon A$ so that: $v_1 x = 0$, $v_2 x = 0$, $v_3 x = v_3$, $v_4 x = v_1$. Now $v_1(ax + xa) = v_2 x + 0 = 0$. Thus, since $(ax + xa)^2 \varepsilon Z$ and since $v_1(ax + xa)^2 = 0$, we must have that $(ax + xa)^2 = 0$. However, $v_3(ax + xa) = v_4 x + v_3 a = v_1 + v_4$, so $0 = v_3(ax + xa)^2 = (v_1 + v_4)(ax + xa) = v_2 + v_3 \neq 0$ since v_2 and v_3 are linearly independent over D. Thus we arrived at a contradiction.

So we must assume that if $v_1, v_2 = v_1 a$, and v_3 are linearly independent over D then

$$(1) \qquad v_3 a = v_1\lambda_1 + v_2\lambda_2 + v_3\lambda_3 \text{ with } \lambda_1, \lambda_2, \lambda_3 \in D.$$

Since D is the commuting ring of A and since $a^2 = 1$, this yields on right multiplication by a that

$$(2) \qquad v_3 = v_3 a^2 = v_1\lambda_2 + v_2\lambda_1 + v_3 a\lambda_3.$$

From (1) and (2) it follows that $\lambda_1 = \lambda_2$, $\lambda_3 = 1$. Thus $v_3 a = (v_1 + v_2)\lambda + v_3$. From this we get that for all $v \in V$, $va = (v_1 + v_2)\lambda(v) + v$, where $\lambda(v) \in D$. There is a $y \in A$ so that $(ay + ya)^2 \neq 0$, for otherwise we could again use the Levitzki theorems (on $a + 1$) to reach a contradiction with the simplicity of A. But if $(ay + ya)^2 \neq 0 \in Z$, we may assume that $(ay + ya)^2 = 1$ (otherwise we would extend Z and proceed as we did with a in getting $a^2 = 1$). Now, $v(ya) = (v_1 + v_2)\lambda(vy) + vy$.

$$v(ay) = ((v_1 + v_2)\lambda(v) + v)y = (v_1 + v_2)\lambda(v)y + vy.$$

Adding $v(ya)$ and $v(ay)$ we obtain

$$v(ay + ya) = (v_1 + v_2)\lambda(vy) + (v_1 + v_2)\lambda(v)y.$$

Thus,

$$\begin{aligned} v &= v(ay + ya)^2 \\ &= (v_1 + v_2)(ay + ya)\lambda(vy) + (v_1 + v_2)y(ay + ya)\lambda(v). \end{aligned}$$

Letting $u_1 = (v_1 + v_2)(ay + ya)$ and $u_2 = (v_1 + v_2)y(ay + ya)$, we have shown that u_1 and u_2 constitute a basis of V over D. V is then 2 dimensional over D, contradicting that v_1, v_2 and v_3 were linearly independent over D.

So we have finally come down to this situation: given v, w in V, then v, va and w are linearly dependent over D. We claim that there is a v in V so that v and va are linearly independent over D. For if $va = v\lambda$, $\lambda \in D$, then $v = va^2 = v\lambda a = (va)\lambda = v\lambda^2$, so $\lambda^2 = 1$, hence $\lambda = 1$; and since $a \neq 1$, $v \neq va$ for some $v \in V$. Thus for this v, v and va are linearly independent over D. This, combined with the fact that v, va and w are linearly dependent for all $w \in V$, implies that w is a linear combination over D of v and va. Thus V is two dimensional over D. Since A is a dense ring of transformations on V over D, A must be all 2×2 matrices over D. All that remains now is to show that $D = Z$.

We have now reduced the whole problem to the following: let A be the set of all 2×2 matrices over a division ring D of characteristic 2; suppose that for some element $a \in A$, $a \notin Z$, a^2 is the unit matrix and $(ax + xa)^2$ is in

Z, the center, for all x in A; from this we want to conclude that $D = Z$, or what is equivalent in this case, namely that D is commutative.

Now, $a = \begin{pmatrix}\alpha & \beta \\ \gamma & \delta\end{pmatrix}$ where α, β, γ, δ, are in D. For any $r, s \,\varepsilon\, D$ we have that

$$\left\{\begin{pmatrix}\alpha & \beta \\ \gamma & \delta\end{pmatrix}\begin{pmatrix}r & s \\ 0 & 0\end{pmatrix}+\begin{pmatrix}r & s \\ 0 & 0\end{pmatrix}\begin{pmatrix}\alpha & \beta \\ \gamma & \delta\end{pmatrix}\right\}^2 \varepsilon\, Z.$$

By carrying out the computations we see that the first element in the second row of this matrix $\{\ \}^2$ is the sum $\gamma r\alpha r + \gamma r^2\alpha + \gamma rs\gamma + \gamma s\gamma r$. Hence this sum is 0 for all $r, s \,\varepsilon\, D$. In particular, if $s = 0$, $\gamma r\alpha r + \gamma r^2\alpha = 0$. Thus the top identity reduces to $\gamma rs\gamma + \gamma s\gamma r = 0$ for all $r, s \,\varepsilon\, D$. If $\gamma \neq 0$, then $rs\gamma = s\gamma r$, and since $s\gamma$ runs through all the elements of D as s does, we have $rt = tr$ for all $r, t \,\varepsilon\, D$, whence D is commutative. So we suppose that $\gamma = 0$. Using the element $\begin{pmatrix}r & 0 \\ s & 0\end{pmatrix}$ of A and proceeding as before, we obtain that $\beta = 0$. It follows that a^2 is the unit matrix. From this we have $\alpha^2 = \delta^2 = 1$; since we are in a division ring of characteristic 2, this means that $\alpha = \delta = 1$, and so a is the unit matrix forcing it to lie in Z, contrary to supposition. In this way we have proved that D is commutative.

Thus A has now become revealed as nothing but a total matrix algebra of degree 2 over a commutative field of characteristic 2. In the course of the proof we have extended the center Z on several occasions, possibly changing A somewhat, but under these various extensions *the dimensions of the new rings over their centers have not changed.* Thus we have shown that Theorem 3 is correct even in characteristic 2, provided A is not 4-dimensional over its center. This then leads us to our principal theorem:

THEOREM 5. *If A is a simple ring and if U is a Lie ideal of A, then either U is contained in the center of A, or else U contains $[A, A]$, except in the case that A is a 4-dimensional simple algebra over its center, which is a field of characteristic 2.*

THE UNIVERSITY OF PENNSYLVANIA.

REFERENCES.

[1] J. Levitzki, "A problem of A. Kurosh," *Bulletin of the American Mathematical Society*, vol. 52 (1946), pp. 1033-1035.

[2] ———, "On the radical of a general ring," *ibid.*, vol. 49 (1943), pp. 462-466.

SOLUTION OF SOME PROBLEMS OF DIVISION.*[1]

Part II. Division by a Punctual Distribution.

By LEON EHRENPREIS.

In Part I of this series [3] it was shown that every constant-coefficient partial differential equation possesses an elementary solution and, moreover, if D is any polynomial of derivation and S a distribution of finite order, then the equation

$$(1) \qquad D * T = S$$

has a solution T which is a distribution of finite order. Moreover, if S is an indefinitely differentiable function, we can choose for T an indefinitely differentiable function. In this paper, we shall extend the above results to the case in which D is any punctual distribution (that is, linear combination of derivatives of measures which represent masses concentrated at a single point). Thus, the results for differential equations are extended to differential-difference equations with constant coefficents. In particular, every differential-difference equation with constant coefficients has an elementary solution.

The idea of the proof is to use the space $\boldsymbol{D}_F$ (see [3]) which, with a slightly different topology, is the Fourier transform of the space $\mathfrak{D}$ of L. Schwartz (see [7]). Under Fourier transformation, the distribution D goes over into an exponential polynomial P. We then must show that, if $PF_\alpha \to 0$ in $\boldsymbol{D}_F$, then also $F_\alpha \to 0$ in $\boldsymbol{D}_F$. This results from a certain minimum modulus property of P which we shall explain now in case we are dealing with distributions on the real line. Then P is an exponential polynomial of one variable, and has a certain "periodic" nature (see [6]): There exists an $L > 0$ so that, in every interval of the real axis of length L, there is a point at which the modulus of P does not get too small. Also, P is an entire function of exponential type. Thus, by using the minimum modulus results of [5], we can produce certain circles on which the modulus of P does not get too small. Once this is established, the result follows as in [3] for pure polynomials by means of the maximum modulus principle.

All the results of this paper can be extended to systems of equations as in [3].

---

* Received August 31, 1954.

[1] Work partially supported by National Science Foundation grant NSF 5-G1010.

The notation of this paper will be that of Part I of this series (see [3], Section 2), and the results of that paper will be used here in an essential way.

2. Solution of the problem. By an exponential polynomial we shall mean here a linear combination of products of polynomials by pure imaginary exponentials. Given any punctual distribution E, there is an exponential polynomial $Q = \boldsymbol{F}E$ such that, for any $f \in \mathcal{D}$, $\mathfrak{F}(E * f) = QF$ (see [4], [7]); conversely, any exponential polynomial Q may be written as $\boldsymbol{F}E$ for some punctual distribution E.

Let Q be an exponential polynomial different from zero one one variable, and write $Q = Q_1 \exp(ia_1 \cdot) + \cdots + Q_k \exp(ia_k \cdot)$, where the Q_j are polynomials $\neq 0$ and where $a_1 < a_2 < \cdots < a_k$. Call $L = a_k - a_1$ and let $Q_1 = b_0 X^m + b_1 X^{m-1} + \cdots + b_m$. (It is allowed that $b_0 = 0$.)

LEMMA 1. *In the above notation, there exists a $K > 0$ which depends only on $a_1, a_2, \cdots, a_k$ and on the degrees of Q_j such that each interval I in the complex plane which is parallel to the real axis and has length $\geqq 2(L+1)$ contains a point y at which*

$$|Q(y)| \geqq K\,|b_0| \exp(-|a_1\alpha|), \tag{2}$$

where α is the distance of I from the real axis.

Remark. In case $L = 0$, this is an easy consequence of the results of [3].

Proof. Let c be the midpoint of I, and call $S = \tau_{-c}Q$. Then S is again an exponential polynomial. Moreover, the exponentials that occur in S are the same as those that occur in Q and, for each exponential, the polynomial coefficient is of the same degree in S as in Q. In addition, if b_0' denotes the coefficient of $X^m \exp(ia_1 \cdot)$ in S, then

$$|b_0'| \geqq |b_0| \exp(-|a_1\alpha|). \tag{3}$$

Now, let us use the theory of mean-periodic functions of L. Schwartz (see [6]). We can find a measure u whose carrier is contained in $[-L-1 \leqq x \leqq L+1]$ which depends only on $a_1, a_2, \cdots, a_k$ and the degrees of the polynomial coefficients of $\exp(ia_j \cdot)$ such that $b_0' = \int S\,du$. Let v be some point in $|x| \leqq L+1$ at which S attains its maximum. Then, if K^{-1} denotes the total variation of u,

$$|b_0'| = \left|\int S\,du\right| \leqq K^{-1}|S(v)| \text{ or } |S(v)| \geqq K\,|b_0'| \geqq K\,|b_0|\, exp(-|a_1\alpha|),$$

by (3). The result follows immediately.

LEMMA 2. *With Q as above, suppose that* $|(\tau_x Q)(z)| \leqq M(x)\exp(l\,|z|)$ *for all complex numbers x, z, where M is a positive function and $l > 0$. Let y be any complex number which satisfies* (2). *Then for any $r > 0$ we can describe a circle about y of radius r' such that $r \leqq r' \leqq 2r$ and*

$$\min_{|z|=r'} |Q(z)| \geqq M(y)^B \exp(clr') K^d\, |b_0|^d \exp(-d\,|a_1\alpha|),$$

where B, c, and d are constants and where d is a positive integer.

Proof. This is just Theorem 5 of Section 4 of [5].

Now, let D be a punctual distribution different from zero and let $P = \boldsymbol{F}D$. By $P\boldsymbol{D}_F$ we denote the space of functions PG for $G \varepsilon \boldsymbol{D}_F$ with the topology induced by $\boldsymbol{D}_F$ (see [3]). The space $D * \mathfrak{D}_F$ is defined similarly.

THEOREM 1. *$PG \rightarrow G$ is a continuous linear map of $P\boldsymbol{D}_F \rightarrow \boldsymbol{D}_F$.*

Proof. The proof will be by induction on the number of letters p which occur in P. For $p = 0$, P is a non-zero constant and the result is obvious. Let $p > 0$ and assume that the theorem is proven for all exponential polynomials in fewer than p letters. Without loss in generality, we may assume that we can write

$$\begin{aligned} P(z) = {} & [P_0(z_2,\cdots,z_p)z_1^m + P_1(z_2,\cdots,z_p)z_1^{m-1} + \cdots + P_m(z_2,\cdots,z_p)] \\ & \times \exp(ia_1z_1) + Q_2(z_1,\cdots,z_p)\exp(ia_2z_1) + \cdots + Q_k(z_1,\cdots,z_p)\exp(ia_kz_1), \end{aligned}$$

where the P_j and the Q_j are exponential polynomials (the exponentials occurring in the Q_j being independent of z_1) with $P_0 \neq 0$, and where $a_1 < a_2 < \cdots < a_k$. The proof will be completed if we can show that $PF \rightarrow P_0{}^dF$ is a continuous linear map of $P\boldsymbol{D}_F \rightarrow P_0{}^d\boldsymbol{D}_F$, where d is the constant that appears in Lemma 2.

This map being obviously linear, we need prove continuity only at zero. Let N be a neighborhood of zero in $P_0{}^d\boldsymbol{D}_F$. By Theorem 7 of [3] we can find a finite sequence of polynomials $A_1,\cdots,A_t$ and an increasing sequence of positive numbers b_l, so that N contains the set of $G \varepsilon P_0{}^d\boldsymbol{D}_F$ which satisfy, for all l,

$$\max_{z \varepsilon C^j \cap C_{b_l}} |\exp(ilp_j\cdot z)A_r(z)G(z)| \leqq 1,$$

for $r = 1, 2,\cdots,t$ and $j = 1, 2,\cdots,2^n$.

For any complex number x, let $x^* \varepsilon C$ be the point $(x, 0,\cdots,0)$. Then it is easily verified that $\tau_{x^*}P = \boldsymbol{F}[\exp(ix^*\cdot)D]$, where $\exp(ix^*\cdot)D$ is the punctual distribution defined by $\exp(ix^*\cdot)D\cdot g = D\cdot\exp(ix^*\cdot)g$ (see [4], [7]). From this it follows immediately that we can find positive numbers a, b, β such that, for any complex number x and any $z \varepsilon C$,

$$(4) \qquad |(\tau_{x^*}P)(z)| \leqq a(1 + |x|)^\beta \exp(\gamma(|z_1| + \cdots + |z_n|)).$$

Set $L = a_k - a_1$ and let q be an integer $> d\,|\,a_1\,|$. For each l, set $c_l = b_{l+q} + 4(L+1)$. Let s, T be positive constants to be specified later, and let $A_1', A_2', \cdots, A_{st}'$ be the polynomials $TA_1, \cdots, TA_t$, $TXA_1, \cdots, TXA_t$, $\cdots, TX^sA_1, \cdots, TX^sA_t$.[2] Let N' be the set of $G \varepsilon \boldsymbol{PD}_F$ which satisfy, for each l, v, j,

$$\max_{z \varepsilon C^j \cap C_{c_l}} |\exp(ilp_j \cdot z) A'_v(z)\, G(z)| \leqq 1,$$

so that N' is a neighborhood of zero in $\boldsymbol{PD}_F$. We claim that, for a proper choice of T and s, $PG \varepsilon N'$ implies $P_0{}^d G \varepsilon N$.

Let $G \varepsilon \boldsymbol{D}_F$ be so chosen that $PG \varepsilon N'$; fix j and l and let $(z_1, \cdots, z_n) \varepsilon C^j \cap C_{b_l}$. We assume that $I(z_1) \geqq 0$, the other case being treated similarly. Denote by P' the exponential polynomial in one variable: $z' \to P(z', z_2, \cdots, z_p)$, and $P_0' = P_0(z_2, \cdots, z_p)$. Let I be the interval in the complex plane which is parallel to the real axis, center z_1, length $2(L+1)$. By Lemma 1 we can find a point $z_0 \varepsilon I$ such that $|P'(z_0)| \geqq K\,|\exp(i\,|\,a_1\,|\,z_0)|\,|P_0'|$, where K is independent of $(z_1, z_2, \cdots, z_n)$. By Lemma 2 and the above, we can describe about z_0 a circle of radius r' with $2(L+1) \leqq r' \leqq 4(L+1)$ such that, for all z' on this circle,

$$|P'(z')| \geqq \delta(1 + |z'|)^B\, |P_0'|^d\, |\exp(id\,|\,a_1\,|\,z_0)|, \tag{5}$$

where B, d, and δ are independent of $(z_1, z_2, \cdots, z_n)$.

Now, for any z' on the circle $|z - z_0| = r'$, $(z, z_2, \cdots, z_n)$ lies in $C_{c_{l-q}}$ if $l \geqq q$, or in C_{c_0} if $l < q$. Moreover by (5), for any z' on this circle and any v,

$$\begin{aligned} &|P_0'|^d\, |A_v(z', z_2, \cdots, z_n) G(z', z_2, \cdots, z_n) \exp(ilp_j \cdot (z', z_2, \cdots, z_n))| \\ &\quad \leqq \eta\, |\exp(-id\,|\,a_1\,|\,z_0 + iqz) \exp(i(l-q)p_j \cdot (z', z_2, \cdots, z_n))| \\ &\qquad \times |A_v(z', z_2, \cdots, z_n) P(z', z_2, \cdots, z_n)(1 + |z'|)^{-B} G(z', z_2, \cdots, z_n)|, \end{aligned} \tag{6}$$

where η is a constant independent of $(z_1, z_2, \cdots, z_n)$. Now,

$$|\exp(-id\,|\,a_1\,|\,z_0 + iqz')| \leqq \exp 4(L+1)q,$$

because $q \geqq d\,|\,a_1\,|$. It follows easily from this that we can choose s and T independent of $(z_1, z_2, \cdots, z_n)$ and of G, l, j, so that the right side of (6) is $\leqq 1$. By the maximum modulus theorem, we also have

$$|P_0'|^d\, |A_v(z_1, z_2, \cdots, \mathrm{z_n}) G(z_1, z_2, \cdots, z_n) \exp(ilp_j \cdot (z_1, z_2, \cdots, z_n))| \leqq 1.$$

We have proven that $PG \to P_0{}^d G$ is a continuous linear map of $\boldsymbol{PD}_F \to P_0{}^d \boldsymbol{D}_F$. But, $P_0{}^d$ is an exponential polynomial in fewer than p

---

[2] X is the polynomial $X(z_1, \cdots, z_n) = z_1$.

variables. Thus, by our induction hypothesis, $PG \to G$ is a continuous linear map of $P\boldsymbol{D}_F \to \boldsymbol{D}_F$, which is the desired result.

From the above we have

COROLLARY. *$g \to D * g$ is a topological isomorphism of $\mathfrak{D}_F$ onto $D * \mathfrak{D}_F$.*

THEOREM 2. *For any $S \varepsilon \mathfrak{D}_F'$ there exists a $T \varepsilon \mathfrak{D}_F'$ such that $D * T = S$. In particular, every differential-difference equation with constant coefficients possesses an elementary solution in $\mathfrak{D}_F'$.*

Proof. It is easily seen that $\bar{D}$ is a punctual distribution (see [4]).[3] Now, $\bar{D} * g \to S \cdot g$ is a continuous linear function T on $\bar{D} * \mathfrak{D}_F$, because it is the composition of the continuous linear maps $\bar{D} * g \to g$ and $g \to S \cdot g$. By the Hahn-Banach theorem (see [1]), T can be extended to a continuous linear function on $\mathfrak{D}_F$, that is, to an element $T \varepsilon \mathfrak{D}_F'$. We have, for any $g \varepsilon \mathfrak{D}_F$ (see [4]),

$$D * T \cdot g = T \cdot \bar{D} * g = T \cdot \bar{D} * g = S \cdot g.$$

Thus, $D * T = S$, which is the desired result.

The case $S = \delta$ (Dirac's measure) shows that D has an elementary solution (see [7]).

For any $l > 0$, $P\boldsymbol{D}_l$ is the space of PG for $G \varepsilon \boldsymbol{D}_l$ with the topology induced by $\boldsymbol{D}$; the spaces $D * \mathfrak{D}_l$ are defined similarly. $D * \mathcal{E}'$ is the space of $D * S$ for $S \varepsilon \mathcal{E}'$ with the topology induced by $\mathcal{E}'$; $D * \mathcal{E}$ is defined similarly.

THEOREM 3. *For any $l > 0$, $PG \to G$ is a continuous linear map of $P\boldsymbol{D}_l \to \boldsymbol{D}_l$.*

Proof. It is clear that, for some $m > 0$, $P\boldsymbol{D}_l \subset \boldsymbol{D}_m$. The result now follows from Theorem 1 and the fact that the spaces $P\boldsymbol{D}_j$, $\boldsymbol{D}_j$ have the topologies induced by $\boldsymbol{D}_F$ (see [3]).

COROLLARY. *For any $l > 0$, $g \to D * g$ is a topological isomorphism of $\mathfrak{D}_l$ onto $D * \mathfrak{D}_l$.*

THEOREM 4. *$S \to D * S$ is a topological isomorphism of $\mathcal{E}'$ onto $D * \mathcal{E}'$.*

Proof. The map is one-to-one by the Paley-Wiener theorem (see [4], [7]) and is clearly continuous, linear, and onto. We must therefore verify continuity, at zero, of the inverse. By the results of [4], the topology of $\mathcal{E}'$ can be described as follows: A fundamental system of neighborhoods of zero in $\mathcal{E}'$ consists of those sets N for which we can find a compact set $K \subset \mathcal{D}_F$ and a neighborhood of zero M in $\mathcal{D}_F$ so that N consists of those $S \varepsilon \mathcal{E}'$ for which $S * f \varepsilon M$ for all $f \varepsilon K$.

---

[3] For any $S \varepsilon \mathfrak{D}'$, $\bar{S}$ is the distribution $\bar{S} \cdot f = \overline{S \cdot \bar{f}}$ for any $f \varepsilon \mathfrak{D}$. Our definition of convolution differs slightly from that of Schwartz (see [4]).

Let N be a neighborhood of zero in $\mathcal{E}'$, so we can find a compact set $K \subset \mathcal{D}_F$ and a neighborhood of zero M in $\mathcal{D}_F$, so that N contains the set of $S \varepsilon \mathcal{E}'$ which satisfy $S * f \varepsilon M$ for all $f \varepsilon K$. By the corollary to Theorem 1, $D * M$ is a neighborhood of zero in $D * \mathcal{D}_F$. Let U be the set of all $D * \bar{T} \varepsilon \mathcal{E}'$ for which $(D * \bar{T}) * g \varepsilon D * M$ for all $g \varepsilon K$, so U is a neighborhood of zero in $D * \mathcal{E}'$. We claim that $D * \bar{T} \varepsilon U$ implies $T \varepsilon N$. For, if $D * \bar{T} \varepsilon U$ and $g \varepsilon K$ (see [4]), $D * (T * g) = (D * \bar{T}) * g \varepsilon D * M$. Thus, $T * g \varepsilon M$, which means that $T \varepsilon N$. This completes the proof of Theorem 4.

THEOREM 5. *$f \rightarrow D * f$ is a continuous linear map of $\mathcal{E}$ onto $\mathcal{E}$.*

Proof. The continuity and linearity of the map are clear. Let $g \varepsilon \mathcal{D}$; then $h : \bar{D} * S \rightarrow S \cdot g$ is a continuous linear function $\bar{D} * \mathcal{E}'$, since it is the composition of the continuous linear maps: $\bar{D} * S \rightarrow S$ and $S \rightarrow S \cdot g$. h can be extended to a continuous linear function h on $\mathcal{E}'$ by the Hahn-Banach theorem. Now, $\mathcal{E}$ is reflexive (see [7]), so there exists a $k \varepsilon \mathcal{E}$ with $h \cdot T = T \cdot k$ for all $T \varepsilon \mathcal{E}'$. Thus, for any $S \varepsilon \mathcal{E}'$ (see [4]),

$$S \cdot D * k = \bar{D} * S \cdot k = h \cdot \bar{D} * S = h \cdot \bar{D} * S = S \cdot g.$$

By the Hahn-Banach theorem, $D * k = g$, which is the desired result.

3. General remarks.

The method developed in Section 2 can be used to solve other types of division problems. Let us consider first the case where $n = 1$. Let S be a distribution whose Fourier transform (see [4]) can be repersented as $\mathfrak{e}(F)$ where F is an entire function of finite order which is slowly increasing on the real axis.[4] Now (see [4] or [5]), $f \rightarrow S * f$ does *not* map $\mathcal{D}_F \rightarrow \mathcal{D}_F$ unless S is of compact carrier. But, suppose that F is mean-periodic. Then, for certain spaces $\mathfrak{D}$ of functions whose Fourier transforms are spaces of entire functions, $f \rightarrow S * f$ maps $\mathfrak{D} \rightarrow \mathfrak{D}$ and the analog of the corollary to Theorem 1 can be proven for these spaces by means of the formal development (see [6]) of mean-periodic functions, and by use of the methods of Section 2.

In case $n > 1$, we need additional hypotheses (besides mean-periodicity of F) in order to guarantee that our induction procedure will work. This

---

[4] For G a slowly increasing function, $\mathfrak{e}(G)$ is the element of $\mathbf{D}'$:

$$\mathfrak{e}(G) \cdot H = \int G(x) H(x)\,dx \qquad \text{for any } H \varepsilon \mathbf{D}.$$

is because no analog of the formal development of L. Schwartz is known for $n > 1$. It is, however, not difficult to see what hypotheses on S will be sufficient to make our induction procedure work.

The above method will be discussed at length in a future publication.

THE JOHNS HOPKINS UNIVERSITY AND THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932.

[2] J. Dieudonné and L. Schwartz, "La dualité dans les espaces $(\mathcal{F})$ et $(\mathcal{LF})$," *Annales de l'Institut Fourier* (*Grenoble*), vol. 1 (1950), pp. 61-101.

[3] L. Ehrenpreis, "Solutions of some problems of division, Part I. Division by a polynomial of derivation," *American Journal of Mathematics*, vol. 76 (1954), pp. 883-903.

[4] ———, "Analytic functions and the Fourier transform of distributions," to appear in the *Annals of Mathematics*.

[5] ———, "Mean periodic functions, Part I. Varieties whose annihilator ideals are principal," *American Journal of Mathematics*, vol. 77 (1955), pp. 293-328

[6] L. Schwartz, "Théorie générale des fonctions Moyenne-periodique," *Annals of Mathematics*, vol. 48 (1947), pp. 857-929.

[7] ———, *Theorie des Distributions*, vol. I-II, Paris, 1950-1951.

MEAN PERIODIC FUNCTIONS.*[1]

Part I. Varieties Whose Annihilator Ideals are Principal.

By Leon Ehrenpreis.

---

1. Introduction. Let C be complex Euclidean n-space. By $\mathfrak{H}$ we denote the space of entire functions in n complex variables with the topology of uniform convergence on the compact sets of C; thus, $\mathfrak{H}$ is a complete, reflexive, locally convex, metrizable, topological vector space. We shall say that $f \varepsilon H$ is *mean-periodic* (see [12]) if the linear combinations of its translates are not dense in $\mathfrak{H}$. Denote by $\mathfrak{H}'$ the dual of $\mathfrak{H}$ with the topology of uniform convergence on the bounded sets of $\mathfrak{H}$. (Since the closed bounded sets of $\mathfrak{H}$ are compact, $\mathfrak{H}'$ is also given the topology of uniform convergence on the compact sets of $\mathfrak{H}$.) Then $f \varepsilon \mathfrak{H}$ is mean periodic if and only if there exists an $S \varepsilon \mathfrak{H}'$ such that the convolution $S * f = 0$ (see Proposition 5 of Section 2). Let f be mean periodic; our two fundamental problems are the following: Let V be the closure of the set of linear combinations of translates of f. *Does V contain an exponential polynomial?* (An exponential polynomial is a linear combination of products of exponentials by monomials.) If the answer to the former question is in the affirmative, then *is every $g \varepsilon V$ the limit of the exponential polynomials of V?*

More generally, let W be any non-empty set in $\mathfrak{H}$ with W different from $\{0\}$ which is closed with respect to translations, linear combinations, and the topology of $\mathfrak{H}$. We shall call such a set a *variety.* An example of such is the set V above. Now, we can ask the same two questions for W as were asked for V above. In case $n = 1$, these two questions were answered in the affirmative by L. Schwartz (see [12]) and, moreover, it was shown by him that every variety is the closure of the set of linear combinations of the translates of some $f \varepsilon \mathfrak{H}$. The present paper is consecrated towards giving a partial extension of these results to the case $n > 1$.

Let W be a variety. Denote by W' the set of all $S \varepsilon \mathfrak{H}'$ for which $S * f = 0$ for every $f \varepsilon W$. It is easily seen that this is the same as the set of $S \varepsilon \mathfrak{H}'$ such that $S(f) = 0$ for every $f \varepsilon W$. It is clear that W' is a closed ideal (under convolution) in $\mathfrak{H}'$. W' is called the *annihilator ideal* of W.

---

* Received August 31, 1954; revised January 14, 1955.

[1] Work supported by National Science Foundation Grants NSF5-G205 and NSF5-G1010.

The main result of this paper is the following: Let W be a variety whose annihilator ideal is principal. Then W contains an exponential polynomial and, moreover, each $f \varepsilon W$ is the limit of exponential polynomials of W. Thus, any two varieties with the same exponential polynomials are identical, i. e., a variety is completely determined by its exponential polynomials. Another way of putting this is: Let V be the variety generated by the exponential polynomials of W, that is, V is the closure, in H, of the exponential polynomials of W; then $V = W$.

Naturally, the question arises as to whether the above result can be extended to other spaces instead of $\mathfrak{H}$. It will be seen that the possibility of extension depends on the ideal theory of the Fourier transform of the space in question. In fact, our main problem is the following: Let $\mathcal{K}$ be a topological ring consisting of entire functions. When is every closed principal ideal in $\mathcal{K}$ determined by its local ideals (see [3])? We shall show that the answer is in the affirmative if $\mathcal{K}$ is

1. The ring $\mathfrak{H}$ of entire functions (this was shown in [3]).
2. The ring of entire functions of exponential type.
3. The ring of polynomials.
4. The ring of entire functions of finite order.

Naturally, it is possible to extend the above to other rings.

We shall also present here a simplified version of the solution to the fundamental problems of mean periodic functions in case that $n = 1$ even for varieties whose annihilator ideals are non-principal. (The first solution was given by L. Schwartz in [12].)

In part II of this series we shall consider nonprincipal annihilator ideals and arbitrary n.

The principal results of this paper were obtained in the Spring of 1953 while the author was under contract NSF S-G205 with the National Science Foundation. I have learned from Professor L. Schwartz that Malgrange has since obtained similar results and has extended them to other spaces of distributions of Schwartz (see [10]).

Unless otherwise specified, all cross references in this paper refer to the section in which they are given. If, for example, in Section 3 we refer to Proposition 1, it will be understood that the reference is to Proposition 1 of Section 3.

2. Fourier transform and convolution in $\mathfrak{H}'$. For any $x \varepsilon C$, x_j is the j-th component of x; by $\| x \|$ we mean $| x_1 | + | x_2 | + \cdots + | x_n |$. and we define $|x| = \max(|x_1|, |x_2|, \cdots, (x_n|)$. If $x, y \in C$, then $x \cdot y = x_1y_1 + x_2y_2 + \cdots + x_ny_n$, and $\exp(x \cdot)$ is the function on $C: y \to \exp(x \cdot y)$. If f is an entire function of n complex variables, we shall say that f is of exponential type $\leqq a$ (where $a \geqq 0$) if, for any $\epsilon > 0$,

$$f(x) = O[\exp(a + \epsilon) \| x \|)].$$

The exponential type of f is the g. l. b. of all a which satisfy the above relation. For any $S \in \mathfrak{H}'$, $f \in \mathfrak{H}$, we shall usually write $S \cdot f$ instead of $S(f)$ with a similar notation for other dual spaces.

Let $S \in \mathfrak{H}'$; since the topology of $\mathfrak{H}$ is that induced by the space J of continuous functions on C with the topology of uniform convergence on the compact sets of C, S can be extended to a continuous linear function u on J. Now u is, by definition, a measure of compact carrier on C, so that we may write, for $f \in \mathfrak{H}$, $S \cdot f = \int f\, du$. By the *Fourier transform* of S we mean the function M on C:

$$M(z) = S \cdot \exp(iz \cdot) = \int \exp(iz \cdot t)\, du(t) = (\boldsymbol{F}(S))(z).$$

It follows immediately from the fact that u is of compact carrier that the integral above converges uniformly for z in any compact set of C, so that M is an entire function. A simple argument shows that M is an entire function of exponential type. Since the set $\{\exp(-iz \cdot)\}_{z \in C}$ is dense in $\mathfrak{H}$, it follows that M determines S uniquely. (Of course, there are infinitely many measures which represent S, any two of which differ by a measure which is zero on $\mathfrak{H}$.)

Conversely, let N be a given entire function of exponential type on C. Then we wish to show that N is the Fourier transform of some $S \in \mathfrak{H}'$. In case $n = 1$, this follows from a theorem of Borel (see [12], also [16], p. 286). In case of $n > 1$, the result can be proven in a similar manner, but we shall give a different proof because it is also applicable to other situations. Before giving the proof, we shall need a few preliminary considerations.

Denote by $\boldsymbol{H}'$ the space of Fourier transforms of elements of $\mathfrak{H}'$ with the topology to make $\boldsymbol{F}: \mathfrak{H}' \to \boldsymbol{H}'$ a topological isomorphism. For c any complex number, $\mathfrak{J}(c)$ is the imaginary part of c, $\mathfrak{R}(c)$ is the real part of c. For any $b > 0$, C_b is the set of $z = (z_1, z_2, \cdots, z_n) \in C$ such that, for each j, $|\mathfrak{J}(z_j)| \leqq b$. For $j = 1, 2, \cdots, n$, we denote by Z_j the function on $C: z \to z_j$. For any $S \in \mathfrak{H}'$, and any integers $k_1, k_2, \cdots, k_n$,

$$S_{k_1k_2\ldots k_n} = S \cdot Z_1^{k_1} Z_2^{k_2} \cdots Z_n^{k_n}.$$

Now, if $f \in \mathfrak{H}$, let $f(z) = \sum f_{k_1k_2\ldots k_n} z_1^{k_1} z_2^{k_2} \cdots z_n^{k_n}$ be the Taylor expansion of f at zero. Then it is clear that the series $\sum f_{k_1k_2\ldots k_n}Z_1^{k_1}Z_2^{k_2}\cdots Z_n^{k_n}$ converges to f in $\mathfrak{H}$. Thus,

$$S\cdot f=\sum S_{k_1k_2\ldots k_n}f_{k_1k_2\ldots k_n}, \tag{1}$$

the series being known to be convergent. The particular choice of $f=\exp(iz\cdot)$ yields

$$\begin{aligned}(2)\quad S\cdot\exp(iz\cdot)&=(\boldsymbol{F}(S))(z)\\ &=\sum(i^{k_1+k_2+\cdots+k_n}/k_1!k_2!\cdots k_n!)S_{k_1k_2\ldots k_n}Z_1^{k_1}Z_2^{k_2}\cdots Z_n^{k_n}.\end{aligned}$$

We have proven

PROPOSITION 1. *For any* $S\,\varepsilon\,\mathfrak{H}'$, $(i^{k_1+k_2+\cdots+k_n}/k_1!k_2!\cdots k_n!)S_{k_1k_2\ldots k_n}$ *are the coefficients in the Taylor expansion of* $\boldsymbol{F}(S)$ *at zero. In particular,* $(i^{k_1+k_2+\cdots+k_n}/k_1!k_2!\cdots k_n!)S_{k_1k_2\ldots k_n}$ *are the Taylor coefficients, at zero, of an entire function of exponential type.*

It is not very difficult to show directly that the numbers $S_{k_1k_2\ldots k_n}$ render the right side of equation (1) convergent whenever $f_{k_1k_2\ldots k_n}$ are the Taylor coefficients of an entire function, if and only if $(i^{k_1k_2\cdots k_n}/k_1!k_2!\cdots k_n!)S_{k_1k_2\ldots k_n}$ are the Taylor coefficients of an entire function of exponential type. (This will be discussed in great detail in a forthcoming article by the author on infinite derivatives.) From this it is not difficult to show that every entire function of exponential type is the Fourier transform of some $S\,\varepsilon\,\mathfrak{H}'$. However, we prefer to give still a third proof, since it will help to illuminate the structure of the topology of $\boldsymbol{H}'$.

The following proposition describes the convergence of sequences in $\boldsymbol{H}'$. By a similar method we could describe the bounded sets of $\boldsymbol{H}'$. Since $\mathfrak{H}'$ is known to be bornologic (that is, a subset N of $\mathfrak{H}'$ is a neighborhood of zero if, for any bounded set B in $\mathfrak{H}'$, we can find a $b>0$ so that $bB\subset N$), we have a complete description of the topology of $\boldsymbol{H}'$.

PROPOSITION 2. *Let* $\{f^j\}$ *be a sequence of entire functions of exponential type which are in* $\boldsymbol{H}'$. *A necessary and sufficient condition that* $f^j\to 0$ *in* $\boldsymbol{H}'$ *is that there exist an* $a>0$ *such that*

$$\max_{x\,\varepsilon\,C}|e^{-a\|x\|}f^j(x)|\to 0. \tag{3}$$

Denote by S^j *the inverse Fourier transform of* f^j. *Then, if* (3) *is satisfied, we even have* $S^j\cdot h\to 0$ *uniformly on the set of* h *which satisfy*

$$\max_{|x|\leqq a}|h(x)|\leqq 1. \tag{4}$$

Proof. Suppose first that $f^j \to 0$ in $\boldsymbol{H}'$. It is known[2] that this implies that $S^j \cdot k \to 0$ uniformly for k in some neighborhood of zero. Thus, there exists an $a > 0$ so that $S^j \cdot k \to 0$ uniformly on the set N of k which satisfy $\max_{|x| \leqq a} | k(x) | \leqq 1$. For any $z \varepsilon C$, we clearly have $e^{-a\|z\|} \exp(iz \cdot) \varepsilon N$ from which (3) follows immediately.

For any h satisfying (4), denote by $h_{k_1k_2\ldots k_n}$ the Taylor coefficients of h at zero. Then we may write (see [2])

$$h_{k_1k_2\ldots k_n} = (1/(2\pi i)^n \int_\Gamma (h(z)/z_1^{k_1} z_2^{k_2} \cdots z_n^{k_n})\, dz$$

where we may take for the chain Γ the cartesian product of the n circles, center origin, radius a, which lie in each of the coordinate plans. Thus,

$$(5) \qquad | h_{k_1k_2\ldots k_n} | \leqq a^{-(k_1+k_2+\cdots+k_n)}.$$

From (3) it follows easily that, if $f^j{}_{k_1k_2\ldots k_n}$ denote the Taylor coefficients of f^j, then, given $\epsilon > 0$, we can find a J so large that, for $j > J$ and all $k_1, k_2, \cdots, k_n$,

$$f^j{}_{k_1k_2\ldots k_n} \leqq \epsilon a^{k_1+k_2+\cdots+k_n}/k_1!k_2!\cdots k_n!.$$

Thus, by Proposition 1, for any h satisfying (3),

$$\begin{aligned} | S^j \cdot h | &= | \sum f^j{}_{k_1k_2\ldots k_n} h_{k_1k_2\ldots k_n} | \\ &\leqq \epsilon \sum 1/k_1!k_2!\cdots k_n! = \epsilon e^n. \end{aligned}$$

This means that $S^j \cdot h \to 0$ uniformly for all h satisfying (4); since this set of h is clearly a neighborhood of zero in $\mathfrak{H}$, we have the desired result.

From Proposition 2 we deduce immediately

PROPOSITION 3. *Let $\{f^j\}$ be a sequence in $\boldsymbol{H}'$ and a, b and $\delta < a$ be positive numbers such that, for all j,*

$$(6) \qquad | f^j(x) | \leqq b \exp[(a - \delta) \| x \|] \quad \textit{for any } x \varepsilon C.$$

Suppose moreover that $f^j \to 0$ in the topology of $\mathfrak{H}$. Then also $f^j \to 0$ in the topology of $\boldsymbol{H}'$. Moreover, if S^j denotes the inverse Fourier transform of f^j, then $S^j \cdot h \to 0$ uniformly for all h which satisfy (4).

PROPOSITION 4. *The topology of $\boldsymbol{H}'$ is stronger than that induced by $\mathfrak{H}$.*

Proof. Let B be a filter base in $\mathfrak{H}'$ which converges to 0; let K be a compact set in C. For any $f \varepsilon \mathfrak{H}$, denote by $B \cdot f$ the filter base obtained from B by replacing each S of B by $S \cdot f$. Now, it is clear that $\{\exp(iz \cdot)\}_{z \varepsilon K}$

---

[2] See, for example, [14], vol. I, p. 91, where a similar result is proven for the space $\mathcal{E}$ of Schwartz.

is a bounded set in H. Thus, the filter bases $B\cdot\exp(iz\cdot)$ for $z \varepsilon K$ converge uniformly to zero. The result follows from the definition of the Fourier transform.

THEOREM. *$\boldsymbol{H}'$ consists of all entire functions of exponential type. An entire function f is of exponential type $\leqq c$ if and only if, for any $\epsilon > 0$, the inverse Fourier transform S of f can be represented by a measure u whose carrier is contained in $C^{c+\epsilon}$.*

(For any $d > 0$, C^d is the set of all $z \varepsilon C$ such that $|z| \leqq d$.)

Proof. The sufficiency of the condition being easily verified, we pass to the necessity. Denote by K the vector space of all entire functions of exponential type, so that $\boldsymbol{H}' \subset K$; we want to show that $\boldsymbol{H}' = K$. Now, we know that every polynomial is in $\boldsymbol{H}'$; we shall show that, for any $f \varepsilon K$, we can find a sequence of polynomials f^j which converge to f in the topology of $\mathfrak{H}$ and such that $\{f^j\}$ is a Cauchy sequence in $\boldsymbol{H}'$. Since the space $\boldsymbol{H}'$ is known to be complete, we shall deduce that $f \varepsilon \boldsymbol{H}'$.

To this end, let $f_{k_1k_2\ldots k_n}$ be the Taylor coefficients of f at the origin. Define

$$f^j = \sum_{k_1+k_2+\ldots+k_n \leqq j} f_{k_1k_2\ldots k_n} Z_1^{k_1} Z_2^{k_2} \cdot\cdot\cdot Z_n^{k_n}.$$

Then it is clear that $f^j \to f$ in the topology of $\mathfrak{H}$. Moreover, it follows from Proposition 3 that $\{f^j\}$ is a Cauchy sequence in $\boldsymbol{H}'$. Since $\boldsymbol{H}'$ is complete, we can find a $g \varepsilon \boldsymbol{H}'$ such that $g = \lim f^j$, the limit being taken in $\boldsymbol{H}'$. Now, by Proposition 4, $g = \lim\limits_{\mathcal{H}} f^j$, from which it follows that $g = f$.

It follows also from Proposition 3 that, for any $\epsilon > 0$, S is bounded on the set of $h \varepsilon \mathfrak{H}$ which satisfy

$$\max_{|x| \leqq c+\epsilon} |h(x)| \leqq 1.$$

Thus, S can be represented by a measure of carrier contained in $C^{c+\epsilon}$ which is the desired result.

For any $f \varepsilon \mathfrak{H}$, $z \varepsilon C$, we denote by $\tau_z f$ the function on C: $(\tau_z f)(x) = f(x - z)$; it is clear that $\tau_z f \varepsilon \mathfrak{H}$. $\tau_z f$ is called the *translate* of f by z. We see easily that $(z, f) \to \tau_z f$ is a continuous map of $C \times \mathfrak{H}$ into $\mathfrak{H}$ and also that, for $f \varepsilon \mathfrak{H}$, $z \to \tau_z f$ is an analytic map of C into $\mathfrak{H}$ over every region in the complex plane (see [8]).

Now, for any $f \varepsilon \mathfrak{H}$, $S \varepsilon \mathfrak{H}'$, we define the convolution $S * f$ by $(S * f)(z) = S \cdot \tau_{-z} f$ for any $z \varepsilon C$. If u is any measure which represents S, then we have

$$(S * f)(z) = \int f(x + z)\, du(x).$$

From the integral expression above, it is clear that $S * f \varepsilon \mathfrak{H}$.

Proposition 5. *For any* $S \varepsilon \mathcal{H}'$, $f \varepsilon \mathcal{H}$, $S * f$ *is the limit in* $\mathcal{H}$ *of linear combinations of translates of* f.

Proof. Let u be a measure of compact carrier representing S and let K be a compact set in C. Then, by the continuity of translation, $\{\tau_{-z}f\}_{z \varepsilon K}$ is a bounded set in $\mathcal{H}$; this set of functions is thus equicontinuous on every compact set of C, hence on a compact set $U \subset C$ whose interior contains the carrier of u. It follows from the theory of the Riemann-Stieltjes integral that, given any $\epsilon > 0$, we can find points $x_1, x_2, \cdots, x_r$ in U and complex numbers $u_1, u_2, \cdots, u_r$ such that, for all $z \varepsilon K$,

$$\left| \int \tau_{-z} f \, du - \sum u_j (\tau_{-z} f)(x_j) \right| \leqq \epsilon$$

or, what is the same thing,

$$\left| (S * f)(z) - \sum u_j (\tau_{-x_j} f)(z) \right| \leqq \epsilon,$$

from which the result follows.

Proposition 6. *If* B *is any bounded set in* $\mathcal{H}$, *then* $S \rightarrow S * f$ *are, for* $f \varepsilon B$, *equicontinuous linear maps of* $\mathcal{H}' \rightarrow \mathcal{H}$. *If* K *is any bounded set in* $\mathcal{H}'$, *then* $f \rightarrow S * f$ *are, for* $S \varepsilon K$, *equicontinuous linear maps of* $\mathcal{H} \rightarrow \mathcal{H}$.

Proof. Since the maps are both clearly linear, we need verify equicontinuity only at zero. Let N be a neighborhood of zero in H; we can find positive numbers a, b such that N contains the set of $f \varepsilon H$ for which $\max_{|x| \leqq a} |f(x)| \leqq b$. By the continuity of translation, the set $\{\tau_{-z} f\}_{|z| \leqq a, f \varepsilon B} = B'$ is bounded in $\mathcal{H}$. Let M be the neighborhood of zero in $\mathcal{H}'$ consisting of all $S \varepsilon \mathcal{H}'$, such that $|S \cdot g| \leqq b$ for all $g \varepsilon B'$. Then clearly, for $S \varepsilon M$, $f \varepsilon B$, we have $S * f \varepsilon N$.

On the other hand, since H is metrizable, hence bornologic (see [6]), we can find a neighborhood of zero P in H such that $|S \cdot f| \leqq b$ for all $S \varepsilon K$, $f \varepsilon P$. Let Q be a neighborhood of zero in $\mathcal{H}$ such that, for all $f \varepsilon Q$, $|z| \leqq a$, we have $\tau_{-z} f \varepsilon P$. Then, for $S \varepsilon K$, $f \varepsilon Q$ we clearly have $S * f \varepsilon N$.

For any S, $T \varepsilon \mathcal{H}'$, the convolution $S * T \varepsilon \mathcal{H}'$ is defined by $S * T \cdot f = S \cdot T * f$ for any $f \varepsilon \mathcal{H}$. ($S * T$ is in $\mathcal{H}'$ by Proposition 6.)

Proposition 7. $S, T \rightarrow S * T$ *is a continuous bilinear map of* $\mathcal{H}' \times \mathcal{H}' \rightarrow \mathcal{H}'$. *Moreover,* $S * T = T * S$.

Proof. The map being obviously bilinear, we need verify continuity only at zero. Let N be a neighborhood of zero in $\mathcal{H}'$ and B a bounded set in $\mathcal{H}'$. Then, there is a bounded set $K \subset \mathcal{H}$ so that N contains the set of all

$S \varepsilon \mathfrak{H}'$, such that $|S * f| \leqq 1$ for all $f \varepsilon K$. By Proposition 5, the set $U = \{T * f\}_{T \varepsilon B, f \varepsilon K}$ is bounded in $\mathfrak{H}$. Thus, there is a neighborhood M of zero in $\mathfrak{H}'$ such that, for $S \varepsilon M$, $g \varepsilon U$, $|S \cdot g| \leqq 1$. Hence, for $S \varepsilon M$, $T \varepsilon B$, $S * T \varepsilon N$.

Now, let u be a measure of compact carrier representing S, and v a measure of compact carrier representing T. Then, for any $f \varepsilon H$,

$$\begin{aligned} S * T \cdot f = S \cdot T * f &= \int du(z) \int f(x+z)\, dv(x) \\ &= \int dv(x) \int f(x+z)\, du(z) = T \cdot S * f = T * S \cdot f. \end{aligned}$$

Thus, $T * S = S * T$.

From the above it follows that, if V is any bounded set in $\mathfrak{H}'$, then $T \to S * T$ are, for $S \varepsilon V$, equicontinuous maps of $\mathfrak{H}' \to \mathfrak{H}'$. Thus, in the terminology of J. Dieudonné and L. Schwartz (see [6], [7]), $S, T \to S * T$ is a separately continuous bilinear map of $\mathfrak{H}' \times \mathfrak{H}' \to \mathfrak{H}$. Since H is a complete metrizable space (hence of type $(\mathcal{F})$), this map must be continuous.

Now, let us describe $S * f$ in terms of the Taylor coefficients of f and $\boldsymbol{F}(S)$ at zero. Let u be a measure of compact carrier which represents S. Then we see easily that, if t denotes the map of $C \to H : t(x) = \tau_{-x} f$, then $\int t\, du$ defines the function $S * f$ of $\mathfrak{H}$. It follows that, if D is any operator of partial differeniation of the form $D = \partial^{k_1+k_2+\cdots+k_n}/\partial Z_1^{k_1} \partial Z_2^{k_2} \cdots \partial Z_n^{k_n}$, then we have

$$(D(S * f))(0) = \int (Df)\, du. \tag{7}$$

We have thus

PROPOSITION 8. *For any integers $k_1, k_2, \cdots, k_n$, and any $S \varepsilon \mathfrak{H}'$, $f \varepsilon \mathfrak{H}$, we have*

$$(S * f)_{k_1 k_2 \ldots k_n} = \sum S_{l_1 l_2 \ldots l_n} f_{l_1+k_1, l_2+k_2, \ldots, l_n k_n}, \tag{8}$$

the sum on the right being extended over all integers $l_1, l_2, \cdots, l_n$, the series on the right being absolutely convergent.

From this we deduce

THEOREM 2. *For any $S, T \varepsilon \mathfrak{H}'$,*

$$\boldsymbol{F}(S * T) = \boldsymbol{F}(S) \boldsymbol{F}(T). \tag{9}$$

Proof. From (8) we have, for any $f \varepsilon \mathfrak{H}$,

$$\begin{aligned} S * T \cdot f = S \cdot T * f &= \sum S_{k_1 k_2 \ldots k_n} (T * f)_{k_1 k_2 \ldots k_n} \\ &= \sum_{k_1 k_2, \cdots, k_n} S_{k_1 k_2 \ldots k_n} \sum_{l_1 l_2, \cdots, l_n} T_{l_1 l_2 \ldots l_n} f_{l_1+k_1, l_2+k_2, \ldots, l_n+k_n}. \end{aligned} \tag{10}$$

Now, $S_{k_1k_2\dots k_n}/k_1!k_2!\cdots k_n!$ and $T_{l_1l_2\dots l_n}/l_1!l_2!\cdots l_n$ are the Tayor coefficients, at the origin, of entire functions of exponential type. Thus, we can find an $a > 1$ so that

$$S_{k_1k_2\dots k_n} = O(a^{k_1+k_2+\cdots+k_n}), \text{ and } T_{l_1l_2\dots l_n} = O(a^{l_1 l_2+\cdots+l_n}).$$

Since f is an entire function,

$$f_{k_1k_2\dots k_n} = O((na)^{-(k_1+k_2+\cdots+k_n)}).$$

Using these estimates we have, for some $\alpha > 0$,

$$\begin{aligned} &| f_{k_1+l_1,+k_2+l_2,\dots,k_n+l_n} S_{k_1k_2\dots k_n} T_{l_1l_2\dots l_n} | \\ &\leqq \sum \alpha (na)^{-(k_1+k_2+\cdots+k_n)-(l_1+l_2+\cdots+l_n)} a^{k_1+k_2+\cdots+k_n} a^{l_1+l_2+\cdots+l_n} \\ &= \alpha \sum n^{-(k_1+k_2+\cdots+k_n)-(l_1+l_2+\cdots+l_n)}. \end{aligned}$$

Since the number of lattice points in real $2n$ dimensional Euclidean space in the circle of center origin, radius r is $O(r^{2n})$, the series above clearly converges. Thus, we may interchange the order of summation in (10) to obtain

$$\begin{aligned} S * T \cdot f &= \sum_{k_1k_2\cdots k_n} S_{k_1k_2\dots k_n} \sum_{l_1l_2\cdots l_n} T_{l_1-k_1,l_2-k_2,\dots,l_n-k_n} f_{l_1l_2\dots l_n} \\ &= \sum_{l_1l_2\cdots l_n} f_{l_1l_2\dots l_n} \sum_{k_1k_2\cdots k_n} S_{k_1k_2\dots k_n} T_{l_1-k_1,l_2-k_2,\dots,l_n-k_n}. \end{aligned}$$

We conclude that the coefficients of the Taylor expansion of $\boldsymbol{F}(S * T)$ at the origin are

$$\begin{aligned} \boldsymbol{F}(S * T)_{l_1l_2\dots l_n} &= S_{k_1k_2\dots k_n} T_{l_1-k_1,l_2-k_2,\dots,l_n-k_n} \\ &= \boldsymbol{F}(S)\boldsymbol{F}(T)_{l_1l_2\dots l_n}. \end{aligned}$$

Thus, $\boldsymbol{F}(S * T) = \boldsymbol{F}(S)\boldsymbol{F}(T)$, which is the desired result.

Remark. Theorem 2 could be easily proven by direct use of the definitions, but we have preferred the above proof because it illustrates the relationship between the Fourier transform and the Taylor coefficients of entire functions. This relationship will be made the object of a detailed study by the author in a forthcoming paper on infinite derivatives.

From this we have immediately

COROLLARY. *For S, T, $U \varepsilon \mathfrak{H}'$, $S * (T * U) = (S * T) * U$, that is, convolution is associative.*

Let us denote by $\boldsymbol{H}$ the dual of $\boldsymbol{H}'$, since $\mathfrak{H}$ is a reflexive, $\mathfrak{H}$ is topologically isomorphic to $\boldsymbol{H}$, the topological isomorphism which we denote by $\boldsymbol{F}$ being given by

$$\boldsymbol{F}(f)\cdot\boldsymbol{F}(S)=S\cdot f,$$

for any $f\,\varepsilon\,\mathfrak{H}$, $S\,\varepsilon\,\mathfrak{H}'$. For any $a\,\varepsilon\,C$ and any non-negative integers $k_1, k_2, \cdots, k_n$, we denote by $\delta_a^{k_1k_2\cdots k_n}$ the element of $\boldsymbol{H}$

$$\delta_a^{k_1+k_2+\cdots+k_n}f=i^{k_1+k_2+\cdots+k_n}[(\partial^{k_1+k_2+\cdots+k_n}/\partial Z_1^{k_1}\partial Z_2^{k_2}\cdots\partial Z_n^{k_n})f](a), \tag{11}$$

for any $f\,\varepsilon\,\mathfrak{H}'$.

PROPOSITION 9. *For any $a\,\varepsilon\,C$ and any integers $k_1, k_2, \cdots, k_n$, we have*

$$\boldsymbol{F}(Z_1^{k_1}Z_2^{k_2}\cdots Z_n^{k_n}\exp(ia\cdot))=\delta_a^{k_1k_2\cdots k_n}. \tag{12}$$

Proof. For any $S\,\varepsilon\,\mathfrak{H}'$, write $f=\boldsymbol{F}(S)$; then

$$\begin{aligned} S\cdot Z_1^{k_1}Z_2^{k_2}\cdots Z_n^{k_n}\exp(ia\cdot) \\ &=S\cdot\sum[(ia)^{l_1+l_2+\cdots+l_n}/l_1!l_2!\cdots l_n!]Z_1^{k_1+l_1}Z_2^{k_2+l_2}\cdots Z_n^{k_n+l_n} \\ &=\sum[(ia)^{l_1+l_2+\cdots+l_n}/l_1!l_2!\cdots l_n!]S_{k_1+l_1,k_2+l_2,\ldots,k_n+l_n} \\ &=\delta_a^{k_1k_2\cdots k_n}\cdot f, \end{aligned}$$

by Taylor's formula (see [2]), where for $a=(a_1, a_2, \cdots, a_n)$ we have written $(ia)^{l_1+l_2+\cdots+l_n}$ for $i^{l_1+\cdots+l_n}a_1^{l_1}a_2^{l_2}\cdots a_n^{l_n}$.

We shall sometimes consider the formula (11) for $f\,\varepsilon\,\mathfrak{H}$ as defining the element $\delta_a^{k_1k_2\cdots k_n}\,\varepsilon\,\mathfrak{H}'$.

3. The spectrum of a variety. By the results of Section 2, we may consider $\mathfrak{H}$ as a topological module, under convolution, over the commutative topological ring $\mathfrak{H}'$. By Proposition 5 of Section 2 and the fact that, for $f\,\varepsilon\,\mathfrak{H}$, $a\,\varepsilon\,C$, $\delta_a * f=\tau_{-a}f$, we see that the varieties of $\mathfrak{H}$ are exactly the closed submodules of $\mathfrak{H}$ which are stable under $\mathfrak{H}'$. Moreover, by Proposition 6 of Section 2, the annihilator ideal of any variety is closed.

By an *ideal* in $\mathfrak{H}'$ ($\boldsymbol{H}'$) we shall mean a closed ideal different from $\mathfrak{H}'$ ($\boldsymbol{H}'$). Let I be an ideal in $\mathfrak{H}'$ and consider the set W of $f\,\varepsilon\,\mathfrak{H}$ for which $S\cdot f=0$ for all $S\,\varepsilon\,I$. We claim that W is a variety. By the Hahn-Banach theorem, W is not reduced to $\{0\}$. Let $f\,\varepsilon\,W$; we claim that, for any $T\,\varepsilon\,\mathfrak{H}'$, $T*f$ is again in W. For, if $S\,\varepsilon\,I$, then

$$S\cdot T*f=S*T\cdot f=0,$$

because $S*T\,\varepsilon\,I$. Thus, W is stable under $\mathfrak{H}'$; since W is closed by Proposition 6 of Section 2, W is a variety. We call W the *annihilator variety* of I.

PROPOSITION 1. *Let W be a variety and W' its annihilator ideal. Then the annihilator variety of W' is W. Conversely, if I is an ideal and I' its annihilator variety, then I is the annihilator ideal of I'.*

This is an immediate result of a general theorem on bipolars in topological vector spaces (see [6]).

Proposition 1 shows us that the correspondence between varieties and their annihilator ideals is a one-one correspondence between the varieties of $\mathfrak{H}$ and the ideals of $\mathfrak{H}'$. Under Fourier transform, this yields a one-one correspondence between the varieties of $\mathfrak{H}$ and the ideals in the ring $\boldsymbol{H}'$ of entire functions of exponential type. Thus, any problem involving the varieties of H can be translated into a problem in the theory of ideals in rings of analytic functions.

Our main task now is to translate the fundamental problems of mean periodic functions into problems about the ring $\boldsymbol{H}'$. Let us note the following first: Let W be a variety and $Z_1^{k_1}Z_2^{k_2}\cdots Z_n^{k_n}\exp(a\cdot)\in W$. Then, since derivation is the limit of linear combinations of translations, it follows that $Z_1^{l_1}Z_2^{l_2}\cdots Z_n^{l_n}\exp(a\cdot)\in W$ whenever $l_1\leqq k_1, l_2\leqq k_2,\cdots, l_n\leqq k_n$. Now, let us denote by $\{a_j\}$ a fixed ordering of the 2^n-1 sequences $j_1, j_2,\cdots, j_r$ for $1\leqq j_1<j_2<\cdots<j_r\leqq n$ and any $r\leqq n$. Then, for any $j=1,2,\cdots,2^n-1$, let $a_j=j_1,j_2,\cdots,j_r$ and let $z\in C$; we call s_j the j-th *part of the spectrum of W at z*, where s_j is the set of r-tuples $(l_1,l_2,\cdots,l_r)$, such that

$$Z_{j_1}{}^{l_1}Z_{j_2}{}^{l_2}\cdots Z_{j_r}{}^{l_r}\exp(z\cdot)\in W$$

but, for no r-tuple $(k_1,k_2,\cdots,k_r)$ with $k_1\geqq l_1, k_2\geqq l_2,\cdots,k_r\geqq l_r$ and some $k_p>l_p$, is $Z_{j_1}{}^{k_1}Z_{j_2}{}^{k_2}\cdots Z_{j_r}{}^{k_r}\exp(z\cdot)\in W$. In some cases, we may have to allow some of the l_p in some of the r-tuples of s_j to be infinity, for example, the r-tuple $(\infty,\infty,l_3,l_4,\cdots,l_r)$ will be part of s_j if

$$Z_{j_1}{}^{l_1}Z_{j_2}{}^{l_2}\cdots Z_{j_r}{}^{l_r}\exp(z\cdot)\in W$$

for any integers l_1, l_2, but no $r-2$ tuples $(k_3,k_4,\cdots,k_r)$ with

$$k_3\geqq l_3, k_4\geqq l_4,\cdots,k_r\geqq l_r$$

which is different from $(l_3,l_4,\cdots,l_r)$ can be used in place of $l_3,l_4,\cdots,l_r$. By the *spectrum of W* we mean the set of 2^n-tuples $(a,s_1,s_2,\cdots,s_{2^n-1})$ where $\exp(a\cdot)\in W$ and s_j is the j-th part of the spectrum of W at a.

Thus defined, the spectrum of W is indeed a very complicated affair, since each s_j consists of a finite, but unbounded, number of r-tuples of integers. In case that $n=1$, it is clear that s_1 (which is the only s_j that appears) is just a single integer. Thus, in this case, our spectrum coincides with the spectrum as defined by L. Schwartz [12].

Now, let I be an ideal in $\boldsymbol{H}'$. By the *cospectrum of* I we mean the spectrum of the annihilator variety of the inverse Fourier transform of I. By Proposition 9 of Section 2, we may describe the cospectrum of I as follows: The cospectrum of I consists of all 2^n tuples $(a, s_1, s_2, \cdots, s_{2^n-1})$ such that a is a common zero of all the functions of I and, for each j, if a_j corresponds to the sequence $j_1, j_2, \cdots, j_r$ then s_j is the set of r-tuples $(l_1, l_2, \cdots, l_r)$ such that, for any $f \varepsilon I$,

$$[(\partial^{l_1+l_2+\cdots+l_r}/\partial Z_{j_1}{}^{l_1}\partial Z_{j_2}{}^{l_2}\cdots\partial Z_{j_r}{}^{l_r})f](a) = 0, \tag{1}$$

but for no other r-tuple $(k_1, k_2, \cdots, k_r)$ with $k_1 \geqq l_1, k_2 \geqq l_2, \cdots, k_r \geqq l_r$ is equation (1) satisfied for all $f \varepsilon I$. In some cases, we may have to allow some of the l_p to be infinity; the definition in that case is analogous to the definition of the spectrum when infinity appeared in some of the r-tuples.

From our previous remarks it follows that if $(l_1, l_2, \cdots, l_r)$ is such that equation (1) is satisfied for all $f \varepsilon I$, then if $k_1 \leqq l_1, k_2 \leqq l_2, \cdots, k_r \leqq l_r$, equation (1) is satisfied for all $f \varepsilon I$ with $(k_1, k_2, \cdots, k_r)$ replacing $(l_1, l_2, \cdots, l_r)$.

For any two varieties V, W, we shall say that the spectrum of V is contained in the spectrum of W if every exponential polynomial of V is in W. For two ideals I, J in $\boldsymbol{H}'$, we say that the cospectrum of I is contained in the cospectrum of J if the spectrum of the annihilator variety of the inverse Fourier transform of I is contained in the spectrum of the annihilator variety of $\boldsymbol{F}^{-1}(J)$. This is the same as saying:

If $(z, s_1, s_2, \cdots, s_{2^n-1})$ is in the cospectrum of I, then for any j, if $a_j = j_1, j_2, \cdots, j_r$ and if $(l_1, l_2, \cdots, l_r)$ is an r-tuple of s_j, then equation (1) holds for all $f \varepsilon J$. (With a suitable modification if some $l_p = \infty$.)

For any $f \varepsilon \boldsymbol{H}'$ and any ideal I of $\boldsymbol{H}'$, we denote by fI the closure of the set of fg for $g \varepsilon I$, so fI is again an ideal. We shall show later that, in fact, $\{fg\}_{g \varepsilon I}$ is itself closed. If J is another ideal in $\boldsymbol{H}'$, we write IJ for the closure of $\{fg\}_{f \varepsilon I,\, g \varepsilon J}$.

Now, let V, W be two varieties. Then clearly, $V \cap W$ is again a variety, or $\{0\}$. By $V + W$ we denote the smallest variety containing V and W (this is easily seen to be the closure of the set of $f + g$ for $f \varepsilon V, g \varepsilon W$). We have the obvious

PROPOSITION 2. *Let V, W be varieties, and V', W' their respective annihilator ideals. Then the annihilator ideal of $V + W$ is $V' \cap W'$, and the annihilator ideal of $V \cap W$ is $V' + W'$ (the smallest ideal containing V' and W') or $\mathfrak{H}'$ if $V \cap W = \{0\}$.*

We shall now introduce three operations on the spectra (or cospectra). Let V and W be varieties and M, P their respective spectra.

(a) The *spectral intersection* $M \cap P$ consists of all 2^n-tuples $(z, s_1, s_2, \ldots, s_{2^n-1})$ such that there exist 2^n-tuples $(z, m_1, m_2, \cdots, m_{2^n-1}) \varepsilon M$ and $(z, p_1, p_2, \cdots, p_{2^n-1}) \varepsilon P$, with the property that if $(l_1, l_2, \cdots, l_r)$ is an r-tuple of s_j, then there exist r-tuples $(k_1, k_2, \cdots, k_r)$ of m_j and $(q_1, q_2, \cdots, q_r)$ of p_j such that $l_1 \leqq k_1, l_2 \leqq k_2, \cdots, l_r \leqq k_r, l_1 \leqq q_1, l_2 \leqq q_2, \cdots, l_r \leqq q_r$. In addition, if $(l_1, l_2, \cdots, l_r)$ is in s_j, then there is no other r-stuple $(t_1, t_2, \cdots, t_r)$ in s_j with $t_1 \leqq l_1, t_2 \leqq l_2, \cdots, t_r \leqq l_r$. (Of course, suitable modifications have to be made with infinity.)

(b) The *spectral sum* $M + P$ consists of all 2^n-tuples $(z, s_1, s_2, \cdots, s_{2^n-1})$ such that there exists either a 2^n-tuple $(z, m_1, m_2, \cdots, m_{2^n-1}) \varepsilon M$ or a 2^n-tuple $(z, p_1, p_2, \cdots, p_{2^n-1})$ with the properties that

1. If there is no 2^n-tuple $(z, m_1, m_2, \cdots, m_{2^n-1}) \varepsilon M$, then $s_j = p_j$ for all j.

2. If there exists no 2^n-tuple $(z, p_1, p_2, \cdots, p_{2^n-1}) \varepsilon P$, then $s_j = m_j$ for all j.

3. If both 2^n-tuples exist, then for each j we form s_j' as the collection of all r-tuples that belong to both p_j and m_j. In this case s_j is the largest subset of s_j' that fulfills the condition that, if $(l_1, l_2, \cdots, l_r)$ is in s_j *and* if $(k_1, k_2, \cdots, k_r)$ is another r-tuple with $k_1 \leqq l_1, k_2 \leqq l_2, \cdots, k_r \leqq l_r$, then $(k_1, k_2, \cdots, k_r) \varepsilon s_j$.

(c) The *spectral product* MP consists of those 2^n-tuples $(z, s_1, s_2, \cdots, s_{2^n-1})$ for which there exists a 2^n-tuple

$$(z, m_1, m_2, \cdots, m_{2^n-1}) \varepsilon M \quad \text{or} \quad (z, p_1, p_2, \cdots, p_{2^n-1}) \varepsilon P$$

such that

1. If $(z, m_1, m_2, \cdots, m_{2^n-1})$ does not exist, then $s_j = p_j$ for all j.

2. If $(z, p_1, p_2, \cdots, p_{2^n-1})$ does not exist, then $s_j = m_j$ for all j.

3. If both exist then, for each j we form the set s_j' consisting of all r-tuples of the form $(k_1 + t_1, k_2 + t_2, \cdots, k_r + t_r)$, where $(k_1, k_2, \cdots, k_r) \varepsilon m_j$ and $(t_1, t_2, \cdots, t_r) \varepsilon p_j$. Here s_j is obtained from s'_j exactly as in the definition of the spectral sum.

If I, J are ideals in $\boldsymbol{H}'$ and O, Q their respective cospectra, then the *cospectral intersection* $O \cap Q$, the *cospectral sum* $O + Q$, and the cospectral product OQ are defined in exactly the same manner.

PROPOSITION 3. *Let V and W be varieties and M, P their respective spectra. Then the spectrum of $V \cap W$ is just $M \cap P$. Let I, J be ideals in*

$\boldsymbol{H}'$ and O, Q their cospectra. Then the cospectrum of $I+J$ is $O\cap Q$ and the cospectrum of IJ is OQ.

Proof. The first part is obvious and the second part then follows by Proposition 2. The third part follows easily from the definitions and the product formula for differentiation.

The question naturally arises as to whether the spectrum of $V+W$ is $M+P$, or, what is the same thing, if the cospectrum of $I\cap J$ is $O+Q$. In case that $n=1$, this question has been answered in the affirmative by L. Schwartz (see [12]). We shall discuss the general case presently.

From Proposition 3 we deduce immediately

COROLLARY. *For any ideals I, J in $\boldsymbol{H}'$, the cospectrum of IJ contains that of I. In particular, if $f\,\varepsilon\,\boldsymbol{H}'$, then the cospectrum of fI contains that of I.*

We shall now formulate the main problems of the theory of mean-periodic functions:

Problem 1. Does every variety contain an exponential? Or, does every ideal in $\boldsymbol{H}'$ have a zero?

Problem 2. Is every variety determined by its exponential polynomials, that is, if W is a variety and $f\,\varepsilon\,W$, if f the limit of exponential polynomials of W? Equivalently, if I is an ideal in $\boldsymbol{H}'$ and if $f\,\varepsilon\,\boldsymbol{H}'$ has the property that the cospectrum of f contains that of I, is $f\,\varepsilon\,I$? (The cospectrum of a $g\,\varepsilon\,\boldsymbol{H}'$ is the cospectrum of the principal ideal generated by g.)

Problem 3. If V, W are varieties and M, P their respective spectra, is the spectrum of $V+W$ equal to $M+P$? That is, if I and J are ideals in $\boldsymbol{H}'$ and if their respective cospectra are N and O, is $N+O$ the cospectrum of $I\cap J$?

Problem 4. What conditions must the spectrum of a variety satisfy?

Problem 5. If W is a variety, does there exist a single $f\,\varepsilon\,W$ such that W is the variety generated by f?

Problem 6. Is every ideal in $\boldsymbol{H}'$ finitely generated?

In case $n=1$, Problems 1, 2, 3, 5, and 6 were answered in the affirmative by L. Schwartz in [12]. The solution to Problem 4 in that case is an immediate consequence of the results of that paper: Let $\{(z_j,n_j)\}=R$ be a sequence of tuples, where the z_j are distinct complex numbers with no limit point and the n_j are integers. Let $\{a_k\}$ be an enumeration of the complex numbers z_j, where each z_j appears n_j times in $\{a_k\}$, and where $|a_k| \leqq |a_{k+1}|$ for all k. A necessary and sufficient condition the R be the spectrum of a variety is that

$$(12) \qquad \limsup |a_k|/k < \infty.$$

For, if R is the cospectrum of an ideal in $\boldsymbol{H'}$, then there is an entire function of exponential type which vanishes at each z_j with order n_j; thus, condition (12) holds. On the other hand, if condition (12) holds, then the function $f(z) = z^p \prod_{a_j \neq 0} (1 - z^2/a_j^2)$ is an entire function of exponential type (where p is the number of $a_j = 0$). R is then the cospectrum of the ideal generated by the functions

$$f_k(z) = z^p (z_k - z)^{n_k} \prod_{a_j \neq 0, a_j \neq z_k} (1 - z^2/a_j^2).$$

For general n, a problem analogous to Problem 4, but for rings of indefinitely differentiable functions instead of $\boldsymbol{H'}$, was discussed by H. Whitney in [18], but even in that case the results are only very sketchy

Assume now that the answer to Problem 2 is in the affirmative, and let W be a variety. Let $\{P_k\}$ denote an enumeration of the monomials in n variables. For each k, consider the set Q_k of points a such that $P_k \exp(a \cdot) \varepsilon W$. We claim that each Q_k is closed in C. For, if I denotes the Fourier transform of the annihilator ideal of W then, by Proposition 9 of Section 2, there is a partial differential operator D_k such that Q_k may be described as the set of points ib where b is a common zero of all functions $D_k g$ for $g \varepsilon I$.

For each k, let $\{c_j^k\}$ be a sequence of points in Q_k whose closure is Q_k. We enumerate all the points c_j^k in some sequence and, for each j, k, we choose a complex number d_j^k in such a manner that the series $\sum d_j^k \exp(c_j^k z) P_k(z)$ converge uniformly for z in any compact set of C and so define an entire function $f \varepsilon W$.

It seems reasonable to exepect that the variety generated by f contains all the functions $\exp(c_j^k \cdot) P_k$, so that, by our assumption, the variety generated by f would be W. However, we are not able to demonstrate this fact, which would imply that every variety is generated by a single element (Problem 5).

Concerning Problem 3, it is clear that $M + P$ is contained in the spectrum of $V + W$. Call I, J the Fourier Transforms of the respective annihilator ideals of V, W. By Proposition 1 the Fourier transform of the annihilator ideal of $V + W$ is $I \cap J \supset IJ$. Now IJ is clearly not reduced to zero so that $V + W \neq \mathfrak{A}$.

Assume now that the solution to Problem 1 is in the affirmative. Then the exponential polynomials of $V+W$ are just the limits of the linear combinations of the exponential polynomials of V and W. We do not know whether the solution to Problem 3 is in the affirmative even in this case.

In what follows we shall solve Problem 2 in case that I is a principal ideal. We shall also give a more simple solution to Problems 1-6 in case that $n=1$ than that given by Schwartz in [12].

4. Solution of the main problem. In order to solve the main problem (Problem 2), we shall make use of the following extension of the Weierstrass preparation theorem (see [2]):

THEOREM 1. *Let P be a function of the complex variables $w, z_1, z_2, \cdots, z_k, x_1, x_2, \cdots, x_r$, which is analytic at the origin, let*

$$D=\partial^{p_1+p_2+\cdots+p_k}/\partial z_1^{p_1}\partial z_2^{p_2}\cdots\partial z_k^{p_k}$$

be a partial differential operator and let s be a positive integer with the following properties:

1. $P(0)=0$.

2. *If E is any partial differential operator in the letters $w, z_1, \cdots, z_k$, but for which at least one of these variables is omitted, then $(EP)(0)=0$.*

3. *The numbers $p_1, p_2, \cdots, p_k, p_{k+1}=s$ are defined as follows:*

(a) *If $q_1<p_1$ and $q_2, q_3, \cdots, q_{k+1}$ are any integers, then*

$$[(\partial^{q_1+q_2+\cdots+q_{k+1}}/\partial z_1^{q_1}\partial z_2^{q_2}\cdots\partial z_k^{q_k}\partial w^{q_{k+1}})P](0)=0.$$

(This defines p_1 uniquely.)

(b) *Suppose $p_1, p_2, \cdots, p_j$ have been defined. Then p_{j+1} is defined by the condition that, if $q_{j+1}<p_{j+1}$ and $q_{j+2}, \cdots, q_{k+1}$ are any integers, then*

$$[(\partial^{p_1+p_2+\cdots+p_j+q_{j+1}+\cdots+q_{k+1}}/\partial z_1^{p_1}\partial z_2^{p_2}\cdots\partial z_j^{p_j}\partial z_{j+1}^{q_{j+1}}\cdots\partial z_k^{q_k}\partial w^{q_{k+1}})P](0)=0.$$

(c) $$(\partial^s DP/\partial w^s)(0)=1.$$

(Thus, the numbers p_j are all defined.)

Then for any function B of $w, z_1, \cdots, z_k, x_1, \cdots, x_r$ which is analytic at the origin there is a unique function Q, analytic at the origin, such that $D(QP)-B$ is, when considered as a power series in w with coefficients functions of $z_1, \cdots, z_k, x_1, \cdots, x_r$ which are analytic at the origin, a polynomial of degree $<s$.

Remark 1. It is allowed that $k=0$, or $r=0$. In case $k=0$, our theorem reduces to the classical Weierstrass preparation theorem.

Remark 2. If P is any function which is analytic at the origin and satisfies $P(0)=0$, then we can find an operator D such that the hypotheses of Theorem 1 are satisfied. For, let P be an analytic function of the complex variables $t_1, t_2, \cdots, t_{r+k+1}$. Call $k+1$ the smallest integer such that there exists a partial differential operator E involving $k+1$ of the variables (say $t_1, t_2, \cdots, t_{k+1}$) with $(EP)(0)\neq 0$. Call $t_1=z_1$, $t_2=z_2, \cdots, t_k=z_k$, $t_{k+1}=w$, $t_{k+2}=x_1, \cdots, t_{r+k+1}=x_r$; the integers $p_1, p_2, \cdots, p_k, s$ are then defined by part 3 of the hypotheses of Theorem 1.

Proof of Theorem 1. Write $P=\sum P_u w^u$ where P_u are analytic functions, at zero, of z_j, x_j. For each u, let

$$P_u=\sum P_{uv_1v_2\ldots v_kv_{k+1}} z_1^{v_1} z_2^{v_2}\cdots z_k^{v_k},$$

where $P_{uv_1v_2\ldots v_kv_{k+1}}$ is the homogeneous term of order v_{k+1} in the expansion of P_u as a power series in the x_j

By the product rule for differentiation, we may write

$$D(QP)=PDQ+\cdots+QDP.$$

We denote the general term in the above sum by $P''Q'$, and we write

$$P''=\sum P''_u w^u, \qquad P''_u=\sum P''_{uv_1\ldots v_{k+1}} z_1^{v_1}\cdots z_k^{v_k},$$

with a similar notation for Q', B, etc.

We are trying to satisfy the equations

$$[D(QP)]_m - B_m = 0 \text{ for } m \geqq s,$$

or, what is the same thing,

$$[D(QP)]_{mv_1\ldots v_{k+1}} - B_{mv_1\ldots v_{k+1}} = 0,$$

whenever $m\geqq s$ and $v_1, v_2, \cdots, v_{k+1}\geqq 0$. Now, $[D(QP)]_{mv_1\ldots v_{k+1}}$ is a sum of terms of the form

$$(Q'P'')_{mv_1\ldots v_{k+1}} = \sum_{u=0}^{m} \sum_{n_1=0}^{v_1} \cdots \sum_{n_{k+2}=0}^{v_{k+1}} Q'_{un_1\ldots n_{k+1}} P''_{m-u, v_1-n_1, \ldots, v_{k+1}-n_{k+1}}.$$

From our hypotheses, we get

(a) For $P''\neq DP$, $P''_{u0\ldots 0}=0$ for all u.

(b) $DP_{u0\ldots 0}=0$ for $u<s$ and $DP_{s0\ldots 0}=1$.

Thus,

$$[D(QP)]_{mn_1n_2\ldots n_{k+1}} = A + (QDP)_{mn_1\ldots n_{k+1}},$$

where

$$A = \sum_{P'' \neq DP} \sum_{i=1}^{k+1} \sum_{u=0}^{m} \sum_{v_1=0}^{n_1} \cdots \sum_{v_i=0}^{n_i-1} \cdots \sum_{v_{k+1}=0}^{n_{k+1}} (Q'_{uv_1\ldots v_{k+1}} P''_{m-u, n_1-v_1, \ldots, n_{k+1}-v_{k+1}}).$$

(We shall understand that $\sum_{j=0}^{-1} a_j = 0$.) Now, we may write

$$(QDP)_{mn_1\ldots n_{k+1}} = E + \sum_{u=0}^{m} Q_{un_1\ldots n_{k+1}} DP_{m-u,0\ldots0},$$

where

$$E = \sum_{i=1}^{k+1} \sum_{u=0}^{m} \sum_{v_1=0}^{n_1} \cdots \sum_{v_i=0}^{n_i-1} \cdots \sum_{v_{k+1}=0}^{n_{k+1}} (Q_{uv_1\ldots v_{k+1}} DP_{m-u, n_1-v_1, \ldots, n_{k+1}-v_{k+1}}).$$

We have

$$\sum_{u=0}^{m} Q_{un_1\ldots n_{k+1}} DP_{m-u,0\ldots0} = \sum_{u=0}^{m-s-1} Q_{un_1\ldots n_{k+1}} DP_{m-u,0\ldots0} + Q_{m-s,n_1,\ldots,n_{k+1}}$$

$$= G + Q_{m-s,n_1\ldots n_{k+1}} \text{ say.}$$

Thus, in order to satisfy the equations

$$[D(QP)]_{mn_1\ldots n_{k+1}} - B_{mn_1\ldots n_{k+1}} = 0$$

for $m \geqq s, n_1, n_2, \cdots, n_{k+1} \geqq 0$, we must have

$$Q_{m-s,n_1,\ldots,n_{k+1}} = B_{mn_1\ldots n_{k+1}} - A - E - G.$$

Or, replacing m by $m+s$ in the above, the equations to be satisfied are

$$(1) \qquad Q_{mn_1\ldots n_{k+1}} = B_{m+s,n_1\ldots n_{k+1}} - A' - E' - G',$$

where A' is obtained from A by replacing m by $m+s$, and E', G' are obtained similarly from E, G respectively.

For any $m, n_1, \cdots, n_{k+1}$, we define

$$(2) \qquad g(m, n_1, \cdots, n_{k+1}) = m + \sum_{i=1}^{k} c_i n_{k-i+1} + c_{k+1} n_{k+1},$$

where the c_j are defined by

$$c_1 = 1 + s;$$

suppose $c_1, \cdots, c_{j-1}$ have been defined, then

$$c_j = 1 + s + \sum_{i=1}^{j-1} c_i p_{k+1-i}.$$

The proof that equations (1) are solvable will be by induction on $g(m, n_1, \cdots, n_{k+1})$. If $g(m, n_1, \cdots, n_{k+1}) = 0$ then $m = 0$ and each $n_j = 0$. Thus, the terms A' and E' are both zero since they involve sums of the form $\sum_{j=1}^{-1}$. Thus, $Q_{0\ldots0} = B_{s0\ldots0}$.

Assume now that we have proven that equations (1) determine $Q_{uv_1\ldots v_{k+1}}$ uniquely whenever $g(u, v_1, \cdots, v_{k+1}) < g(m, n_1, \cdots, n_{k+1})$. Then, clearly, G' is uniquely determined. On the terms of E', g attains its maximum when, for some $j \leqq k+1$,

$$u = m + s, v_1 = n_1, \cdots, v_{j-1} = n_{j-1}, v_j = n_j - 1, v_{j+1} = n_{j+1}, \cdots, v_{k+1} = n_{k+1}.$$

In that case

$$g(u, v_1, \cdots, v_{k+1}) = m + s + \sum_{i \neq j} c_i n_{k-i+1} + c_{k+1} n_{k+1} + c_j n_{k-j+1} - c_j$$

$$= g(m, n_1, \cdots, n_{k+1}) + s - c_j < g(m, n_1, \cdots, n_{k+1}),$$

because each $c_j \geqq 1 + s$. Thus, E' is also uniquely determined.

Let us write, for $P'' \neq DP$, $Q' = (\partial^{t_1 + \cdots + t_k} / \partial z_1^{t_1} \cdots \partial z_k^{t_k}) Q$. Then, by the hypotheses of the theorem, $P''_{uv_1\ldots v_k 0} = 0$ if

(a) $v_1 < t_1$ and any $u, v_2, \cdots, v_k$.

(b) $v_1 = t_1, \cdots, v_j = t_j, v_{j+1} < t_{j+1}$, and any $u, v_{j+2}, \cdots, v_k$.

(c) $v_1 = t_1, \cdots, v_k = t_k, u < s$.

Consider the terms $Q'_{uv_1\ldots v_{k+1}}$ in A' (for a fixed P'') for which

(α) $u = m + s, n_1 - v_1 = t_1 + 1, v_2 = n_2, \cdots, v_{k+1} = n_{k+1}$.

(β) $u = m + S, \quad n_1 - v_1 = t_1, \cdots, n_j - v_j = t_j, \quad n_{j+1} - v_{j+1} = t_{j+1} + 1,$
$v_{j+2} = n_{j+2}, \cdots, v_{k+1} = n_{k+1}$.

(γ) $u = m + s, v_1 = n_1, \cdots, v_k = n_k, v_{k+1} = n_{k+1} - 1$.

(δ) $u = m, n_1 - v_1 = t_1, \cdots, n_k - v_k = t_k, v_{k+1} = n_{k+1}$.

(It may happen that some of these terms do not occur, but we shall neglect this fact since it will be seen that, in this case, the equations (1) are certainly solvable.)

Given any term $Q'_{u'v'_1\ldots v'_{k+1}}$ in A' for this same P'', let us notice that this term is uniquely determined by $Q_{u', v'_1+t_1, \ldots, v'_k+t_k, v'_{k+1}}$. From (a), (b), and (c), given any such term for which the subscripts $u', v'_1, \cdots, v'_{k+1}$ do not satisfy any of the conditions (α), (β), (γ), or (δ), we can find a term $Q'_{uv_1\ldots v_{k+1}}$ whose subscripts satisfy one of these conditions, and such that

$$g(u', v'_1 + t_1, \cdots, v'_k + t_k, v'_{k+1}) < g(u, v_1 + t_1, \cdots, v_k + t_k, v_{k+1}).$$

We shall show that in cases (α), (β), and (γ)

$$g(u, v_1 + t_1, \cdots, v_k + t_k, v_{k+1}) < g(m, n_1, \cdots, n_{k+1}).$$

Case (α). We have

$$\begin{aligned} g(u, v_1+t_1, \cdots, v_k+t_k, v_{k+1}) &\leqq m + s\sum_{i=1}^{k-1} c_i(n_{k-i+1}+p_{k-i+1}) \\ &\qquad + c_k(n_1-1) + c_{k+1}n_{k+1} \\ &= m + \sum_{i=1}^{k} c_i n_{k-i+1} + c_{k+1}n_{k+1} + s + \sum_{i=1}^{k-1} c_i p_{k-i+1} - c_k \\ &= g(m, n_1, \cdots, n_{k+1}) - 1. \end{aligned}$$

Case (β).

$$\begin{aligned} g(u, v_1+t_1, \cdots, v_k+t_k, v_{k+1}) &\leqq m + s + c_{k+1}n_{k+1} \\ &\quad + \sum_{i=1}^{k-j-1} c_i(n_{k-i+1}+p_{k-i+1}) + c_{k-j}(n_{j+1}-1) + \sum_{i=k-j+1}^{k} c_i n_{k+1-i} \\ &= g(m, n_1, \cdots, n_{k+1}) + s + \sum_{i=1}^{k-j-1} c_i p_{k-i+1} - c_{k-j} \\ &= g(m, n_1, \cdots, n_{k+1}) - 1. \end{aligned}$$

Case (γ).

$$\begin{aligned} g(u, v_1+t_1, \cdots, v_k+t_k, v_{k+1}) &\leqq m + s + \sum_{i=1}^{k} c_i(n_{k-i+1}+p_{k-i+1}) + c_{k+1}(n_{k+1}-1) \\ &= g(m, n_1, \cdots, n_{k+1}) - 1. \end{aligned}$$

Case (δ). In this case, clearly,

$$g(u, v_1+t_1, \cdots, v_k+t_k, v_{k+1}) = g(m, n_1, \cdots, n_{k+1}).$$

The term in question is

$$Q'_{m,n_1-t_1,\ldots,n_k-t_k,n_{k+1}}P''_{s,t_1\ldots t_k0} = dQ_{mn_1\ldots n_{k+1}}P_{sp_1\ldots p_k0} = dQ_{mn_1\ldots n_{k+1}},$$

because $P_{sp_1\ldots p_k0} = 1$, where d is some constant $\geqq 0$ (which may depend on P'').

Thus, equation (1) takes the form

$$(1 + \sum d)\, Q_{mn_1\ldots n_{k+1}} = B_{m+s,n_1\ldots n_{k+1}} + R_{mn_1\ldots n_{k+1}}, \tag{3}$$

where $R_{mn_1\ldots n_{k+1}}$ depends only on those $Q_{uv_1\ldots v_{k+1}}$ for which

$$g(u, v_1, \cdots, v_{k+1}) < g(m, n_1, \cdots, n_{k+2}).$$

Thus, equation (1) determines the $Q_{mn_1\ldots n_{k+1}}$ uniquely.

Analyticity of Q. We have shown that Q exists and is unique as a formal power series; we shall now show that this power series converges in the neighborhood of the origin.

We shall produce positive numbers $k, a, a_1, \cdots, a_{k+1}$ all > 2 so that

$$|\, Q_{mn_1n_2\ldots n_{k+1}}(x) \,| \leqq K a^m a_1{}^{n_1} a_2{}^{n_2} \cdots a_{k+1}{}^{n_{k+1}} \tag{4}$$

for all x in a suitable neighborhood of zero N in complex Euclidean r-space X. Since nothing is altered in either the hypotheses or conclusion of Theorem 1 upon application of the linear transformation

$$w \rightarrow cw, z_1 \rightarrow cz_1, \cdots, z_k \rightarrow cz_k, x_1 \rightarrow cx_1, \cdots, x_r \rightarrow cx_r$$

for $c > 0$, we may assume that both B and P are analytic in the unit ball in complex Euclidean $k + r + 1$-space. (The unit ball in C is the set of $z \varepsilon C$ such that $|z| \leqq 1$.) Thus, there exist constants $L > 0$, $M > 0$ so that, for all $m, n_1, \cdots, n_{k+1}$, and all $x \varepsilon X$ with $|x| \leqq 1$, we have

$$|B_{mn_1\ldots n_{k+1}}(x)| \leqq L, \qquad |P''_{mn_1\ldots n_{k+1}}(x)| \leqq M,$$

for all P''.

The proof that equation (4) is satisfied will be by induction on $g(m, n_1, \cdots, n_{k+1})$. For N we choose the unit ball in X. Next, choose $K > 2$ so that for all $x \varepsilon N$

$$|Q_{00\ldots 0}(x)| \leqq K.$$

(Such a K can be found by virtue of the definition of $Q_{00\ldots 0}$.)

Now, assume (4) is satisfied for all $u, v_1, \cdots, v_{k+1}$ such that

$$g(u, v_1, \cdots, v_{k+1}) < g(m, n_1, \cdots, n_{k+1}) > 0.$$

Set

$$G(m, n_1, \cdots, n_{k+1}) = K a^m a_1^{n_1} \cdots a_{k+1}^{n_{k+1}}.$$

Let us examine the terms $Q'_{uv_1,v_2\ldots v_{k+1}}$ which appear on the right side of equation (3). Each is a constant multiple of some term $Q_{u'v'_1\ldots v'_{k+1}}$ with

$$g(u'v'_1 \cdots v'_{k+1}) < g(m, n_1, \cdots, n_{k+1}).$$

Moreover, each of these constant multipliers does not exceed

$$(n_1 + p_1)^{p_1}(n_2 + p_2)^{p_2} \cdots (n_k + p_k)^{p_k}.$$

On the other hand, there is a term d on the left side of equation (3) with

$$d = n_1(n_1 - 1) \cdots (n_1 - p_1) n_2(n_2 - 1) \cdots (n_2 - p_2) \cdots n_k(n_k - 1)(n_k - p_k),$$

where it is understood that, in the above product, each term which is $\leqq 0$ is to be replaced by 1. Thus, the ratio of the maximum constant multiplier on the right side of equation (3) to $1 + \sum d$ does not exceed

$$h = (2p_1)^{p_1}(2p_2)^{p_2} \cdots (2p_k)^{p_k}.$$

Further, no term $Q_{u'v'_1\ldots v'_{k+1}}$ can be involved on the right side of equation (3) more than

$$q = p_1! \, p_2! \cdots p_k!$$

times.

From the above considerations and equation (3) we have, for any $x \varepsilon N$,

$$\begin{aligned} |\,Q_{mn_1\ldots n_{k+1}}(x)\,| \leqq L + Mhq \times \Big[& \sum_{u=0}^{m+s} \sum_{v_1=0}^{n_1-1} \sum_{v_2=0}^{n_2+p_2} \cdots \sum_{v_k=0}^{n_k+p_k} \sum_{v_{k+1}=0}^{n_{k+1}} G(u, v_1, \cdots, v_{k+1}) \\ & + \sum_{u=0}^{m+s} \sum_{v_2=0}^{n_2-1} \sum_{v_3=0}^{n_3+p_3} \cdots \sum_{v_k=0}^{n_k+p_k} \sum_{v_{k+1}=0}^{n_{k+1}} G(u, n_1, v_2, \cdots, v_{k+1}) \\ & + \cdots + \sum_{u=0}^{m+s} \sum_{v_k=1}^{n_k-1} \sum_{v_{k+1}=0}^{n_{k+1}} G(u, n_1, \cdots, n_{k-1}, v_k, v_{k+1}) \\ & + \sum_{u=0}^{m+s} \sum_{v_{k+1}=0}^{n_{k+1}-1} G(u, n_1, \cdots, n_k, v_{k+1}) + \sum_{u=0}^{m-1} G(u, n_1, \cdots, n_{k+1}) \Big], \end{aligned}$$

because of (α), (β), (γ), and (δ) above.

Thus,

$$\begin{aligned} |\,Q_{mn_1\ldots n_{k+1}}(x)\,| \leqq L + MKqh \times \Big[& \frac{a^m}{a-1} a_1^{n_1} \cdots a_{k+1}^{n_{k+1}} \\ & + a^{m+s+1} a_1^{n_1} \cdots a_k^{n_k} \frac{a_{k+1}^{n_{k+1}}}{a_{k+1}-1} \\ & + a^{m+s+1} a_1^{n_1} \cdots a_{k-1}^{n_{k-1}} \frac{a_k^{n_k}}{a_k-1} a_{k+1}^{n_{k+1}+1} \\ & + a^{m+s+1} a_1^{n_1} \cdots a_{k-2}^{n_{k-2}} \frac{a_{k-1}^{n_{k-1}}}{a_{k-1}-1} a_k^{n_k+p_k+1} a_{k+1}^{k+1+1} \\ & + \cdots + a^{m+s+1} \frac{a_1^{n_1}}{a_1-1} a_2^{n_2+p_2+1} a_k^{n_k+p_k+1} a_{k+1}^{n_{k+1}+1} \Big]. \end{aligned}$$

We want to make this $\leqq K a^m a_1^{n_1} \cdots a_{k+1}^{n_{k+1}}$. This will be the case if

$$\begin{aligned} L + MKhq \times \Big[& \frac{1}{a-1} + a^{s+1} \frac{1}{a_{k+1}-1} + a^{s+1} a_{k+1} \frac{1}{a_k-1} + a^{s+1} a_{k+1} a_k^{p_k+1} \frac{1}{a_{k-1}-1} \\ & + \cdots + a^{s+1} a_{k+1} a_k^{p_k+1} \cdots + a_2^{p_2+1} \frac{1}{a_1-1} \Big] \leqq K. \end{aligned}$$

This can be done by choosing successively $a, a_1, \cdots, a_{k+1}$ all > 2 so that each summand above is $\leqq K/(k+2)(L+MKhq)$.

With this choice of $a, a_1, \cdots, a_{k+1}$ the induction process can be completed. Thus relation (4) is satisfied. This means that Q is analytic in a suitable neighborhood of the origin, which concludes the proof of Theorem 2.

Let $a \varepsilon C$ and let f, g be functions which are analytic in the neighborhood of a. We shall say that the *cospectrum of f contains the cospectrum of g at a* if there exists a neighborhood N of a so that, for any $b \varepsilon N$ and any non-negative integers $t_1, \cdots, t_n$, the conditions

$$[(\partial^{s_1+s_2+\cdots+s_n}/\partial Z_1^{s_1} \partial Z_2^{s_2} \cdots \partial Z_n^{s_n}) g](b) = 0,$$

whenever $s_1 \leqq t_1, s_2 \leqq t_2, \cdots, s_n \leqq t_n$ imply

$$[(\partial^{s_1+s_2+\cdots+s_n}/\partial Z_1^{s_1}\partial Z_2^{s_2}\cdots\partial Z_n^{s_n})f](b) = 0$$

for all $s_1 \leqq t_1, s_2 \leqq t_2, \cdots, s_n \leqq t_n$. In part II of this series we shall discuss in detail the concept of the cospectrum of a local ideal at a (see [3]).

THEOREM 2. *Let $a \varepsilon C$ and let P, g be functions which are analytic in a neighborhood of a. Suppose that the cospectrum of g contains the cospectrum of P at a. Then there is a function h analytic at a such that $g = hP$ in a suitable neighborhood of a.*

Proof. The theorem is trivially true in case that $P = 0$, for then we must have $g = 0$; we therefore assume $P \neq 0$. It is clearly no loss in generality to assume that $a = 0$. If $P(0) \neq 0$, then the result is obvious; we assume also that $P(0) = 0$. $D, w, z_1, \cdots, z_k, x_1, \cdots, x_r$ are now defined as in the hypotheses of Theorem 1 $(r + k + 1 = n)$.

By Theorem 1, there is a unique function S analytic at the origin such that

$$D[P(w,z,x)S(w,z,x)] = w^s + \sum_{j=0}^{s-1} K_j(z,x)w^j, \tag{5}$$

where K_j are analytic at the origin. It is shown in the proof of Theorem 1 that $S(0) = S_{0\ldots 0} = 1$. Thus, the operator D and the function PS also satisfy the hypotheses of Theorem 1. Hence, by Theorem 1, we can find a function Q analytic at the origin and such that

$$D[P(w,z,x)S(w,z,x)Q(w,z,x)] - Dg(w,z,x) = \sum_{j=0}^{s-1} H_j(z,x)w^j, \tag{6}$$

where the H_j are analytic at the origin.

Let M be a ball in C of radius so small that all functions which appear in equations (5) and (6) are analytic in M, and such that $S \neq 0$ in M. Then we integrate equations (5) and (6) p_1 times with respect to z_1, p_2 times with respect to $z_2, \cdots, p_k$ times with respect to z_k, where all paths of integration begin at the origin and lie entirely in M. Then we obtain

$$P(w,z,x)S(w,z,x) = \sum_{j=0}^{s} \tilde{K}_j(z,x)w^j, \tag{7}$$

$$P(w,z,x)S(w,z,x)Q(w,z,x) - g(w,z,x) = \sum_{j=0}^{s-1} \tilde{H}_j(z,x)w^j. \tag{8}$$

Let us note that $K_j(0,0) = 0$ for $j = 0, 1, \cdots, s-1$. For, were this not the case, the integer s would not have required minimum property of the hypotheses of Theorem 1. It is fairly easy to see, moreover, using the properties of the operator D, that, for each $j < s$,

$$\lim_{z\to 0,\ x\to 0} \tilde{K}_j(z,x)/\tilde{K}_s(z,x) = 0.$$

By hypothesis, the cospectrum of g contains the cospectrum of P at 0; since $S(0) \neq 0$, the cospectrum of g contains the cospectrum of PS at 0. Thus, there exists a neighborhood of zero $N \subset M$ so that, for any $b \varepsilon N$ and any integers $t_1, \cdots, t_n$, the conditions

$$[(\partial^{s_1+\cdots+s_n}/\partial Z_1^{s_1}\cdots\partial Z_n^{s_n})PS](b) = 0,$$

for $s_1 \leqq t_1, \cdots, s_n \leqq t_n$, imply

$$[(\partial^{s_1+\cdots+s_n}/\partial Z_1^{s_1}\cdots\partial Z_n^{s_n})g](b) = 0,$$

whenever $s_1 \leqq t_1, \cdots, s_n \leqq t_n$.

Now, for $a_0 \neq 0$, the zeros of the polynomial $a_0X^t + a_1X^{t-1} + \cdots + a_t$ depend continuously on $(a_1/a_0, a_2/a_0, \cdots, a_t/a_0)$. Thus, by the above, we may choose a point (z_0, x_0) such that $K_s(z_0, x_0) \neq 0$ and such that all the zeros c of $\sum_{j=0}^{s} \tilde{K}_j(z_0, x_0) w^j$ have the property that (c, z_0, x_0) lies in N.

We wish to show that all $\tilde{H}_j = 0$. Assume this is not the case; then we may assume that (z_0, x_0) is so chosen that $\tilde{H}_t(z_0, x_0) \neq 0$ for some t.

It follows from the above that if $(b, z_0, x_0) \varepsilon N$ has the property that b is a k-fold zero of $\sum_{j=0}^{s} \tilde{K}_j(z_0, x_0) w^j$, then b is also (at least) a k-fold zero of $\sum_{j=0}^{s-1} \tilde{H}_j(z_0, x_0) w^j$. But, we have seen that $\sum_{j=0}^{s} \tilde{K}_j(z_0, x_0) w^j$ has s zeros $b_1, \cdots, b_s$ counting multiplicity, with $(b_j, z_0, x_0) \varepsilon N$ for all j. Thus $\sum_{j=0}^{s-1} \tilde{H}_j(z_0, x_0) w^j$ has at least s zeros. This can only be the case if all $\tilde{H}_j(z_0, x_0) = 0$. This contradiction concludes the proof of Theorem 2.

We have the immediate

COROLLARY. *Let $a \varepsilon C$, f, g be functions which are analytic at a and such that the cospectrum of f contains the cospectrum of g at a. Let L be a linear transformation of $C \to C$. Then the cospectrum of fL contains the cospectrum of gL at La.*

Remark. If a proof of this corollary could be found which is independent of Theorems 1 and 2, then this corollary, together with the Weierstrass preparation theorem, could be used to give a simplified proof of Theorem 2 without the use of Theorem 1.

Let f, g be entire functions on C. We shall say that *the cospectrum of* f *contains the cospectrum of* g if, for any $a \varepsilon C$, the cospectrum of f contains the cospectrum of g at a. It is clear that, if f and g are of exponential type, this definition coincides with the previous one (see Section 3).

THEOREM 3. *Let* f, g *be entire functions such that the cospectrum of* f *contains the cospectrum of* g. *Then there is an entire function* h *such that* $f = gh$.

Proof. f/g is clearly analytic outside of the set of zeros of g. According to the theorem of removable singularities (see [2]), it is sufficient to prove that f/g is bounded in the neighborhood of each point. This is an immediate consequence of Theorem 2.

The passage from arbitrary entire functions to entire functions of exponential type is accomplished by Theorem 4 below; for $n = 1$, Theorem 4 is a result of E. Lindelöf (see [10]).

THEOREM 4. *Let* f, g, h *be entire functions of* n *complex variables, with* $g = fh \neq 0$. *Suppose that* f *and* g *are of exponential type. Then* h *is also of exponential type.*

The following minimum modulus theorem, which is itself of interest, is the key tool in the proof of Theorem 4:

THEOREM 5. *Let* k *be an entire function of exponential type of one complex variable with* $k(0) \neq 0$ *and* $|k(z)| \leqq M e \exp(A|z|)$, *for all* z. *Then for any* $r > 0$ *there is an* r' *with* $r \leqq r' \leqq 2r$ *such that*

$$\min_{|z|=r'} |k(z)| \geqq \exp(B \log M + cAr' + d \log |k(0)|)$$

where B *and* c *are certain constants, and where* d *is a positive integer.*

Let us assume the truth of Theorem 5 and conclude the proof of Theorem 4. We shall prove by induction on p the following proposition α_p: Let H, K be constants, $f \neq 0$ an entire function of exponential type of p variables, and h an entire function on C $(p \leqq n)$ such that, for any $(z_1, \cdots, z_n) \varepsilon C$,

$$|f(z_1, \cdots, z_p)| \, |h(z_1, \cdots, z_n)| \leqq K \exp(H \| z \|).$$

Then h is of exponential type.

Now, α_0 is trivially true, because for $p = 0$, f is a constant $\neq 0$. Let $p > 0$ and assume α_q is true whenever $q < p$. Suppose that $|f(z'_1, z'_2, \cdots, z'_p)| \leqq L \exp[N(|z'_1| + \cdots + |z'_p|)]$ for all $(z'_1, \cdots, z'_p)$. Let $(z_1, \cdots, z_n)$

be any point in C such that $f(0, z_2, \cdots, z_p) \neq 0$, and also $z_1 \neq 0$. By Theorem 5 we can find an r' with $|z_1| \leqq r' \leqq 2|z_1|$ such that

$$\min_{|z|=r'} |f(z, z_2, \cdots, z_p)| \\ \geqq L^B |f(0, z_2, \cdots, z_p)|^d \exp[BN(|z_2| + \cdots + |z_p|) + cNr'],$$

where B, c, and d are certain constants. (Because $z \to f(z, z_2, \cdots, z_p)$ is an entire function of exponential type of one complex variable which is different from zero at the origin.)

From this it follows that, for any z with $|z| = r'$,

$$L^B |f(0, z_2, \cdots, z_p)|^d \exp(BN(|z_2| + \cdots + |z_p|) + cNr') |h(z, z_2, \cdots, z_n)| \\ \leqq |f(z, z_2, \cdots, z_p)| \, |h(z, z_2, \cdots, z_n)| \\ \leqq Ke^{H(|z|+|z_2|+\cdots+|z_n|)}.$$

Thus,

$$|f(0, z_2, \cdots, z_p)|^d |h(z, z_2, \cdots, z_n)| \\ \leqq KL^{-B} \exp((H - BN)(|z_2| + \cdots + |z_n|) + (H - cN)r') \\ \leqq KL^{-B} \exp((H - BN)(|z_2| + \cdots + |z_n|) + 2(H - cN)|z_n|),$$

because $r' \leqq 2|z|$. Hence, by the maximum modulus theorem

$$|f(0, z_2, \cdots, z_p)|^d |h(z_1, z_2, \cdots, z_n)| \\ \leqq KL^{-B} \exp(H - BN)(|z_2| + \cdots + |z_n|) + 2(H - cN)|z_1|).$$

This inequality holds for all $(z_1, \cdots, z_n)$ such that $z_1 \neq 0$ and $f(0, z_2, \cdots, z_p) \neq 0$; by continuity it holds without these restrictions.

Now, $(z_2, \cdots, z_p) \to f(0, z_2, \cdots, z_p)$ is an entire function of exponential type in $p - 1$ variables. Thus, by our induction hypothesis, h is an entire function of exponential type, i.e., Proposition α_p is established. This completes our induction; Theorem 4 results immediately.

Proof of Theorem 5. We claim it is sufficient to consider the case in which k is an even function. For, assume Theorem 5 is proven for even functions and set $f = k\check{k}$, where $\check{k}(x) = k(-x)$, so f is an even entire function of exponential type. Suppose that $|k(z)| \leqq Me^{A|z|}$, so that $|f(z)| \leqq M^2 e^{2A|z|}$ for all complex numbers z. Then, by the assumed result for even functions, given $r > 0$ we can find an r' with $r \leqq r' \leqq 2r$ such that

$$\min_{|z|=r'} |f(z)| \geqq \exp(2B \log M + 2cA|z| + d \log |f(0)|),$$

for some constants B, c, and d.

Let z_0 be a point on $|z| = r'$ at which $|k|$ attains its minimum. Then

$$|f(z_0)| = |k(z_0)|\,|k(-z_0)| \leqq |k(z_0)| \max_{|z|=r'} |k(z)|.$$

Thus,

$$\begin{aligned} \min_{|z|=r'} |f(z)| &\leqq |k(z_0)| \max_{|z|=r'} |k(z)| \\ &= \min_{|z|=r'} |k(z)| \max_{|z|=r'} |k(z)| \leqq Me^{Ar'} \min_{|z|=r'} |k(z)|. \end{aligned}$$

We conclude that

$$\begin{aligned} \min_{|z|=r'} |k(z)| &\geqq M^{-1}e^{-Ar'} \min_{|z|=r'} |f(z)| \\ &\geqq \exp[(2B - I)\log M + (2c - 1)Ar' + 2d \log|k(0)|], \end{aligned}$$

which is the desired inequality.

Theorem 5 results immediately from the above and

THEOREM 6. *Let k be an entire function of finite order of one complex variable with $k(0) \neq 0$. Call p the genus of k and assume that k is invariant under the linear transformation $z \to \omega z$ of the complex plane into itself, where ω is a primitive $(p+2)$-nd root of unity. Then for any $r > 0$ there is an r' with $r \leqq r' \leqq 2r$ such that*

$$\min_{|z|=r'} |k(z)| \geqq 2\,|k(0)|^{11} \exp\Big[-2 \sum_{j=0}^{\infty} 2^{-j(p+2)} (2^{p+2} - 1) \log M(2^{j+3}r)\Big],$$

where, for any $y > 0$, $M(y) = \max_{|z|=y} |k(z)|$.

Proof of Theorem 6. From the fact that, for any $k < p + 2$, $\sum_{j=1}^{p+2} \omega^{kj} = 0$, we deduce easily that the Weierstrass-Hadamard product for k (see [16]) takes the form $k(z) = \delta \prod (1 - z^{p+2}/a_j^{p+2})$, where $\delta = k(0)$.

We write $k(z) = \delta \pi_1(z) \pi_2(z) \pi_3(z)$ where

$\pi_1(z)$ consists of all terms of $\prod$ for which $|a_j| \leqq r/2$,

$\pi_2(z)$ consists of all terms of $\prod$ for which $r/2 < |a_j| \leqq 4r$,

$\pi_3(z)$ consists of all terms of $\prod$ for which $|a_j| > 4r$.

(By convention, an empty product is to be replaced by unity.)

For π_1 we use the trivial estimate that, for $|z| \geqq r$,

$$|1 - z^{p+2}/a_j^{p+2}| \geqq |z/a_j|^{p+2} - 1 \geqq 2^{p+2} - 1 > 1,$$

so that $|\pi_1(z)| \geqq 1$ whenever $|z| \geqq r$.

For π_2 and π_3 we shall need Jensen's formula (see [16]). If $m(x)$ denotes the number of zeros of k of modulus $\leqq x$, then for any $y > 0$,

$$\int_0^y (m(x)/x)dx = (1/2\pi)\int_0^{2\pi} \log k(ye^{i\theta})d\theta - \log|k(0)| \leqq \log M(y) - \log|\delta|.$$

But, m is a non-decreasing function, so

$$2\int_y^{2y} (m(x)/x)\,dx \geqq m(y)\,2\int_y^{2y} dx/x \geqq m(y).$$

Hence,

$$\tfrac{1}{2}m(y) \leqq \int_0^{2y} (m(x)/x)\,dx \leqq \log M(2y) - \log|\delta|. \tag{10}$$

From (10) we see that in π_2 there are not more than

$$\beta = (2/p+2)(\log M(8r) - \log|\delta|) \text{ terms } a_j.$$

Now (see [11], p. 86, problem 66) we can find an r' such that $r \leqq r' \leqq 2r$ and such that

$$\prod_{a_j \varepsilon \pi_2} |(|a_j| - r')| \geqq 2(r/4)^{\beta}. \tag{11}$$

Also, if $|z| = r'$,

$$\begin{aligned} (12)\quad & |a_j^{p+2} - z^{p+2}| \\ & = (a_j - z)(a_j - \omega z) - \cdots - (a_j - \omega^{p+1}z) \geqq |(|a_j| - r')|^{p+2}. \end{aligned}$$

Thus, if $|z| = r'$,

$$\begin{aligned} \prod_{a_j \varepsilon \pi_2} |1 + z^{p+2}/a_j^{p+2}| &\geqq (4r)^{-\beta(p+2)} \prod_{a_j \varepsilon \pi_2} |a_j^{p+2} - z^{p+2}| \\ &\geqq (4r)^{-\beta(p+2)} \prod_{a_j \varepsilon \pi_2} |(|a_j| - r')|^{p+2} \\ &\geqq (4r)^{-\beta(p+2)} 2(r/4)^{\beta(p+2)} \geqq 2(4)^{-2\beta(p+2)}, \end{aligned}$$

by (11) and (12); that is, $|\pi_2(z)| \geqq 2(4)^{-2\beta(p+2)}$ for $|z| = r'$. It is clear that this implies

$$|\pi_2(z)| \geqq 2|\delta|^8 \exp(-4\log M(8r)). \tag{13}$$

For π_3 we use the estimate

$$|1 - z^{p+2}/a_j^{p+2}| \geqq 1 - |z/a_j|^{p+2} \geqq \exp(-|z/a_j|^{p+2}). \tag{14}$$

This results from the fact that, for $0 < u < \frac{1}{2}$,

$$\log(1-u) = -u - u^2/2 - u^3/3 - \cdots > -u,$$

so that $1 - u > \exp(-u)$.

For any $j \geqq 2$, consider the annulus $R_j\colon 2^j r < z \leqq 2^{j+1}r$. In R_j there are not more than $\beta_j = (2/p+2)(\log M(2^{j+2}r) - \log|\delta|)$ terms a_j by (10). For each of them,

$$|r'/a_j| \leqq 2r/2^j r = 2^{1-j}. \tag{15}$$

Thus, if $|z| = r'$, using (14) and (15),

$$|\pi_3(z)| \geqq \exp(-\sum |r'/a_j|^{p+2}) \geqq \exp(-\sum_{j=2}^{\infty} (p+2)\beta_j 2^{(p+2)(1-j)}).$$

Now, $\sum_{j=2}^{\infty} 2^{(p+2)(1-j)} = 1/2^{p+2} - 1 < 1$. Thus,

$$(16) \qquad |\pi_3(z)| \geqq \exp(-2^{p+2}-1) \sum_{j=2}^{\infty} (p+2)\beta_j 2^{(p+2)(1-j)}$$

$$= |\delta|^2 \exp[-2 \sum_{j=2}^{\infty} 2^{(p+2)(1-j)} (2^{p+2}-1) \log M(2^{j+2}r)],$$

for all z with $|z| = r'$. Putting all our estimates together, we have

$$\min_{z=r'} |k(z)| \geqq |\delta| \min_{z=r'} |\pi_1(z)| \min_{z=r'} |\pi_2(z)| \min_{z=r'} |\pi_3(z)|$$

$$\geqq 2\,|k(0)|^{11} \exp[-\sum_{j=0}^{\infty} 2^{-(p+2)j} (2^{p+2}-1) \log M(2^{j+3}r)],$$

which is the desired result.

Combining Theorems 3 and 4 we have

THEOREM 7. *Every principal ideal in* $\boldsymbol{H}'$ *is completely determined by its cospectrum, that is, if* $f, g \in \boldsymbol{H}'$ *and if the cospectrum of* g *contains the cospectrum of* f, *then* g *lies in the ideal generated by* f *in* $\boldsymbol{H}'$.

THEOREM 8. *Let* $f \in \boldsymbol{H}'$ *and let* $I = \{fh\}_{h \in \boldsymbol{H}'}$. *Then* I *is closed, i.e., every principal algebraic ideal in* $\boldsymbol{H}'$ *is an ideal in* $\boldsymbol{H}'$.

Proof. Let g be in the closure of I; it is clear that the cospectrum of g contains the cospectrum of f. Thus, by Theorems 3 and 4, $g \in I$.

Remark. It is easily seen that the set I above is even closed in $\mathfrak{A}$, and hence weakly in $\boldsymbol{H}'$ (that is, if $\boldsymbol{H}'$ is given the weak topology of the dual of $\boldsymbol{H}'$).

From Theorem 7 and the results of Section 3 we deduce

THEOREM 9. *Let* V *be a variety whose annihilator ideal is principal. Then* V *contains an exponential polynomial. Moreover,* V *is completely determined by its exponential polynomials, that is, every* $f \in V$ *is the limit of exponential polynomials of* V.

Theorem 9 is the generalization to n variables of the theorem of Ritt, Valiron, and others to the effect that every solution of a differential equation of infinite order with constant coefficients is the limit of exponential polynomial solutions of the equation (see [12]).

5. Extension to other spaces. In this section, we shall investigate rings $\mathfrak{O}$ of entire functions with the following property: Let $f, g \varepsilon \mathfrak{O}$ and suppose that the cospectrum of g contains the cospectrum of f. Then there is an $h \varepsilon \mathfrak{O}$ such that $g = fh$. Theorem 2 of Section 4 tells us that, for any $a \varepsilon C$, if $\mathfrak{P}_a$ denotes the ring of functions analytic at a, then every principal ideal in $\mathfrak{P}_a$ is determined by its cospectrum at a (that is, if $F, G \varepsilon \mathfrak{P}_a$ and if the cospectrum of G contains the cospectrum of F at a, then $G = FH$ for some $H \varepsilon \mathfrak{P}_a$). Thus, the above property of $\mathfrak{O}$ can be reformulated as follows: Every principal ideal in $\mathfrak{O}$ is determined by its local ideals, that is, by the ideals it generates in $\mathfrak{P}_a$ for all $a \varepsilon C$ (see [3]). Theorems 3 and 4 of Section 4 show that we may take for $\mathfrak{O}$ the ring of entire functions or the ring of entire functions of exponential type. In this section we shall show that we may take for $\mathfrak{O}$ the ring of polynomials or, for any $c \geqq 0$, the ring of entire functions of order $\leqq c$, hence also the ring of all entire functions of finite order, and more general rings.

THEOREM 1. *Every principal ideal in the ring of polynomials is determined by its local ideals. Equivalently, if P is any polynomial and Q a polynomial whose cospectrum contains the cospectrum of P, then $Q = PS$ for some polynomial S.*

Proof. According to Theorem 3 of Section 4, we must show the following: Let f and g be polynomials in n letters with $f \neq 0$ and let h be an entire function of n complex variables such that $g = fh$; then h is also a polynomial. The proof will be by induction on p of the following proposition: Let $f \neq 0$ be a polynomial in p letters ($p \leqq n$) and h an entire function of n complex variables such that, for all $(z_1, \cdots, z_n) \varepsilon C$,

$$|f(z_1, \cdots, z_p) h(z_1, \cdots, z_n)| \leqq (1 + |z_1| + \cdots + |z_n|)^k,$$

for some $k \geqq 0$. Then h is a polynomial.

For $p = 0$, f is a non-zero constant and the result is immediate. Assume $p > 0$ and that the result is true for polynomials in fewer than p letters. Let $(z_1, \cdots, z_n) \varepsilon C$, and write

$$f(X_1, X_2, \cdots, X_p) = f_0 X_1^q + f_1 X_1^{q-1} + \cdots + f_q,$$

where the f_j are polynomials in $X_2, \cdots, X_p$ and where $f_0 \neq 0$. By the results of [9], we can describe a circle γ of radius $\leqq 1$ about z_1 so that, for all $z \varepsilon \gamma$,

$$|f(z, z_2, \cdots, z_p)| \geqq \beta \, |f_0(z_2, \cdots, z_p)|,$$

where β is a positive constant which depends only on the degree of f. Thus, for any $z \varepsilon \gamma$,

$$\begin{aligned}|\beta f_0(z_2,\cdots,z_p)h(z,z_2,\cdots,z_n)| &\leqq |f(z,z_2,\cdots,z_p)h(z,z_2,\cdots,z_n)| \\ &\leqq (1+|z|+|z_2|+\cdots+|z_n|)^k \\ &\leqq (2+|z_1|+|z_2|+\cdots+|z_n|)^k \\ &\leqq 2^k(1+|z_1|+\cdots+|z_n|)^k.\end{aligned}$$

Hence, by the maximum modulus theorem,

$$|2^{-k}\beta f_0(z_2,\cdots,z_p)h(z_1,z_2,\cdots,z_n)| \leqq (1+|z_1|+\cdots+z_n|)^k.$$

Now, $2^{-k}\beta f_0$ is a polynomial in fewer than p variables. Thus, by our induction assumption, h is a polynomial. Theorem 1 now results immediately.

Let $f \neq 0$ be an entire function of n complex variables. We say that f is of finite order if there is an $A \geqq 0$ so that

$$|f(z)| = O(\exp(\|z\|^A)). \tag{1}$$

The greatest lower bound of all numbers A which satisfy (1) is called the order of f. In order to prove the analog of Theorem 1 for the ring of entire functions of finite order, we shall need the following form of the minimum modulus theorem (see [16])

THEOREM 2. *Let f be an entire function of one complex variable, with $|f(z)| \leqq M\exp(.z|^A)$, and such that $f(0) \neq 0$. Then for any $r > 0$ there is an r' with $r \leqq r' \leqq 2r$ and*

$$\min_{|z|=r'} |f(z)| \geqq \exp(\alpha\log M + \zeta|z|^A + \gamma\log|f(0)|),$$

where α, ζ, and γ are constants which depend only on A, and where we may choose, for γ, a positive integer.

Proof. Call p the genus of f (see [16]). Then, as in the proof of Theorem 5 of Section 4, it is sufficient to consider the case in which f is invariant under the linear transformation $z \rightarrow wz$ of the complex plane into itself, where w is a primitive $(p+2)$-nd root of unity. The result now follows immediately from Theorem 6 of Section 4.

Exactly as in the proof of Theorem 4 of Section 4 (except that we use Theorem 3 instead of Theorem 5 of Section 4), we deduce

THEOREM 4. *For any $A > 0$, denote by $\mathfrak{Q}_A$ the ring of entire functions of order $\leqq A$. If f, $g \varepsilon \mathfrak{Q}_A$ and if h is an entire function such that $fh = g$, then we must have $h \varepsilon \mathfrak{Q}_A$. Every principal ideal in $\mathfrak{Q}_A$ is determined by its local ideals, or, what is the same thing, if $P, S \varepsilon \mathfrak{Q}_A$ have the property that* *the cospectrum of S contains the cospectrum of P, then $S = PT$ for some $T \varepsilon \mathfrak{Q}_A$.*

We have the immediate

COROLLARY. *All the results of Theorem 4 for the rings $\mathfrak{Q}_A$ hold also for the ring $\mathfrak{Q}$ of all entire functions of finite order.*

The results of Theorem 4 and its corollary extend to the following more general situation: Let ϕ be a positive, even, continuous function of one real variable, with ϕ monotonically increasing to infinity. We call ϕ an *admissible* function if

(a) $\phi(x) = O(|x|^p)$ for some p.

(b) There exists a p satisfying (a) such that, for all $r > 0$,

$$\sum_{j=0}^{\infty} 2^{-j(p+2)} \phi(2^{j+3} r) \leqq d\phi(r), \tag{2}$$

where d is a constant depending only on ϕ. A family Φ of admissible functions ϕ is called an *admissible* family if there exists an admissible funtcion ψ (not necessarily in Φ) such that, for every r and every $\phi \varepsilon \Phi$, $\phi(r) \geqq \psi(r)$. As examples of admissible functions we have the functions $\phi(r) = r^p$ for any $p > 0$. An example of an admissible family is all funtcions $\phi(r) = r^p$ for all $p > q \geqq 0$.

If ϕ is an admissible function, we denote by Q_ϕ the ring of entire functions f of n complex variables which satisfy

$$f(z) = O[\exp(\phi(\| z \|))]. \tag{3}$$

For any admissible family Φ, we define $Q_\Phi = \bigcap_{\phi \varepsilon \Phi} Q_\phi$ and, for Ψ an arbitrary family of admissible functions, we define $Q^*_\Psi = \bigcup_{\psi \varepsilon \Psi} Q$. Then we can easily generalize Theorem 4 and its corollary to

THEOREM 5. *Let Φ be an admissible family. Then for any $f, g \varepsilon Q_\Phi$, if h is an entire function satisfying $fh = g$, then we must have $h \varepsilon Q_\Phi$. Every principal ideal in Q_Φ is determined by its local ideals, that is, if $P, S \varepsilon Q_\Phi$ have the property that the cospectrum of S contains the cospectrum of P, then $S = PT$ for some $T \varepsilon Q_\Phi$. These results also hold for the spaces Q^*_Ψ for any family Ψ of admissible functions.*

Remark 1. Theorem 5 obviously contains Theorem 4 and its corollary and Theorem 6 of Section 4 as special cases.

Remark 2. Theorem 5 itself can be extended to more general families, for the condition (b) in the definition of admissible can certainly be improved and even condition (a) can be slightly ameliorated. However, a recent result of Hayman on the minimum modulus makes it extremely unlikely that condition (a) can be entirely omitted (see *Proc. Lond. Math. Soc.*, 1952, pp. 469-512).

6. The case $n=1$. Throughout this section we shall suppose that $n=1$. We shall give a simplification of the solution by L. Schwartz [12] of problem 2 of Section 3. It is fairly simple to prove from this (see [12]) that every ideal in $\boldsymbol{H}'$ is the ideal generated by two functions of $\boldsymbol{H}'$, and that every variety in $\mathfrak{H}$ is the variety generated by a single element. In Section 3 it was shown how this solution of Problem 2 implies that the spectrum of the union of two varieties is the spectral union of their spectra. Thus, the results of this section will conclude the solutions of all the problems of Section 3.

PROPOSITION. *Let I be an ideal in $\boldsymbol{H}'$ and $F \varepsilon I$, $F \neq 0$. Suppose that, for some $a \varepsilon C$, F has a zero of order $m>0$ and some $G \varepsilon I$ has no zero at a. Then $F/(Z-a) \varepsilon I$.*

Proof. Let W be the annihilator variety of the inverse Fourier transform of I; denote by S the inverse Fourier transform of $F/(Z-a)$, and call V the variety generated by $\{S * f\}_{f \varepsilon W}$. From the formula $T \cdot S * f = T * S \cdot f$ for any $f \varepsilon \mathfrak{H}$, $T \varepsilon \mathfrak{H}'$, it follows immediately that T is in the annihilator ideal J of V if and only if $T * S$ is in the annihilator ideal of W.

Call K the Fourier transform of J. The above shows us that $P \varepsilon K$ if and only if $PF/(Z-a) \varepsilon I$. Assume $F/(Z-a) \not\varepsilon I$; then $(Z-a) \varepsilon K$ but $1 \not\varepsilon K$. Thus, V is not $\{0\}$ and, in fact, consists exactly (see Section 2) of the solutions $\alpha \exp(ia\cdot)$ for $\alpha \varepsilon C$ of the ordinary differential equation $-i(df/dZ)-af=0$ for $f \varepsilon \mathfrak{H}$. Hence, $\exp(ia\cdot) \varepsilon W$, which implies (see Section 3) that every $Q \varepsilon I$ has a zero at a. This contradiction completes the proof.

THEOREM 1. *Every variety contains an exponential polynomial, or, what is the same thing, every ideal in $\boldsymbol{H}'$ has a zero.*

Proof. Assume I is an ideal in $\boldsymbol{H}'$ with no zero, and let $f \varepsilon I$, $f \neq 0$. We may assume f is even (for if f is not even, we may consider $f\bar{f} \varepsilon I$). Let

$$f(z) = z^m \prod_{j=1}^{\infty} (1 - z^2/a_j^2)$$

be the Weierstrass-Hadamard factorization of f. For each $k > 0$, set

$$f_k(z) = \prod_{j=k}^{\infty} (1 - z^2/a_j^2).$$

Then by induction on the above proposition, we find easily that each $f_k \varepsilon I$. Moreover, it is easy to see that we can find positive numbers A, M so that, for all k and all $z \varepsilon C$,

$$|f_k(z)| \leqq \prod_{j=1}^{\infty} (1 + |z|^2/|a_j|^2) \leqq Me^{A|z|}.$$

Further, it is clear that $f_k \to 1$ in $\mathfrak{H}$. Thus, by Proposition 3 of Section 2, $f_k \to 1$ in $\boldsymbol{H}'$. Since I is closed, this means that $1 \varepsilon I$ which contradicts the fact that I is an ideal (since an ideal is $\neq \boldsymbol{H}'$).

THEOREM 2. *Every variety is determined by its spectrum. Equivalently, if I is an ideal in $\boldsymbol{H}'$ and if the cospectrum of $f \varepsilon \boldsymbol{H}'$ contains the cospectrum of I, then $f \varepsilon I$.*

Proof. Let V be a variety; call W the variety generated by the exponential polynomials of V and suppose $f \varepsilon V$, $f \not\varepsilon W$. By the Hahn-Banach theorem we can find an $S \varepsilon \mathfrak{H}'$ which is zero on W but such that $S \cdot f \neq 0$. Consider the variety X generated by $\{S * g\}_{g \varepsilon V}$ ($X \neq 0$ because $S * f \neq 0$); we claim X has no exponential polynomials.

Denote by F the Fourier transform of S, and by I, J, K the respective Fourier transforms of the annihilator ideals of V, X, and W. Assume there exists an exponential polynomial in X; then there is an $a \varepsilon C$ at which every $G \varepsilon J$ vanishes. As in the proof of the above proposition, $G \varepsilon J$ if and only if $GF \varepsilon I$. Thus, the above implies that every multiple of F which is in I has a zero at a of order $> m =$ order of zero of F at a. Now, by construction, the cospectrum of F contains the cospectrum of I. Thus, there is a $G \varepsilon I$ whose order k of zero at a is $\leqq m$. Then $(G/(Z-a)^k)F = G(F/(Z-a)^k) \varepsilon I$, but has a zero at a of order m. This contradiction proves that X has no exponential polynomials. This contradicts Theorem 1 and so completes the proof of Theorem 2.

7. General remarks. The methods of Section 6 apply also to other spaces (but still only for $n = 1$), for example the space $\mathcal{E}$ of Schwartz (see [14], [12], and [8]) of indefinitely differentiable functions. In that case, the proof of the proposition of Section 6 can be carried out as before. The proof of Theorem 1 of Section 6 does not seem to extend to $\mathcal{E}$; however, if I is an ideal in $\boldsymbol{E}'$ (see [8]) which has no zero, then for $f \varepsilon I$, $f \neq 0$, and f

rapidly decreasing on the real axis, it is shown by Schwartz (see [12]) that also $f' \varepsilon I$. From this it follows easily (see [12]) that $1 \varepsilon I$. This proves the analog of Theorem 1 of Section 6 for $\boldsymbol{E}$. The analog of Theorem 2 then follows as before.

There are two essential difficulties in extending the methods of Section 6 to $n > 1$. The first is the fact that the ideals in the spaces $\mathfrak{P}_a$ (see Section 5) need not be principal for $n > 1$. Thus, no analog of $F/(Z-a)$ which is used in the proof of the proposition of Section 6 can be found. But even if this difficulty could be surmounted, the following difficulty remains: For $n > 1$, there exist functions in $\mathfrak{A}$ other than exponential polynomials which are solutions of constant-coefficient partial differential equations. Thus, an entirely new approach is needed for the problem. This will be done in Part II of this series.

Each $S \varepsilon \mathfrak{A}'$ defines an element of the space $\boldsymbol{D}'$, or even of $\boldsymbol{E}'$ (see [8]), namely the restriction T of S to $\boldsymbol{D}$. It is clear that $\boldsymbol{D}$ is dense in $\mathfrak{A}$, so that T determines S uniquely. We want to show now that the Fourier transform M of S defined in Section 2 is, essentially, the inverse Fourier transform U of T, as defined in [8]. Using the notation of [8], if $F \varepsilon \boldsymbol{D}$, and if v is a measure of compact carrier which represents S (hence also T), then by an easily justified change in the order of integration,

$$\begin{aligned} U \cdot f = T \cdot f = \int F(z)\,dv(z) &= \int dv(z) \int f(x) e^{-ix\cdot z} dx \\ &= \int f(x)\,dx \int e^{-ix\cdot z} dv(z) = \int M(-x) f(x) = \mathrm{j}\underset{\wedge}{M} \cdot f. \end{aligned}$$

Thus, $U = \mathrm{j}\underset{\wedge}{M}$ which is the desired result.

The Johns Hopkins University and The Institute for Advanced Study.

BIBLIOGRAPHY.

[1] S. Banach, *Théorie des Opérations Linéaires*, Monografje Matematyczne, Warsaw, 1932.

[2] S. Bochner and W. T. Martin, *Several Complex Variables*, Princeton, 1948.

[3] H. Cartan, "Idéaux et modules de fonctions analytiques de n variables complexes," *Bulletin de la Societe Mathematique de France*, vol. 78 (1950), pp. 28-64.

[4] ———, "Variétés analytiques complexes et cohomologie," *Colloque sur les Fonctions de Plusieurs Variables*, Bruxelles, 1953.

[5] – ——, *Faisceaux Analytiques*, Seminaire E.N.S., Paris, 1951-52, Chapt. XV-XX (mimeographed).

[6] J. Dieudonné and L. L. Schwartz, "La dualité dans les espaces ($\mathcal{F}$) et ($\mathcal{LF}$)," *Annales de l'Institut Fourier* (Grenoble), vol. I (1950), pp. 61-101.

[7] L. Ehrenpreis, *Theory of distributions for locally compact spaces*, Columbia University Thesis, 1953 (to appear).

[8] ———, "Analytic functions and the Fourier transform of distributions," *Annals of Mathematics* (to appear).

[9] ———, "Solution of some problems of Division, part I," *American Journal of Mathematics*, vol. 76 (1954), pp. 883-903.

[10] B. Malgrange, "Sur quelques propriétés des equations de convolution," *Comptes Rendus des Séances de l'Académie des Sciences*, vol. 238 (1954), pp. 2219-2221.

[11] G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. II, Berlin, 1925.

[12] L. Schwartz, "Théorie générale des fonctions moyenne-périodiques," *Annals of Mathematics*, vol. 48 (1947), pp. 857-929.

[13] ———, *Etude des Sommes d'Exponentielles Réelles*, Paris, 1943.

[14] ———, *Théorie des Distributions*, vol. I-II, Paris, 1950-51.

[15] ———, "Analyse et synthèse harmonique dan les espaces de distributions," *Canadian Journal of Mathematics*, vol. 3 (1951), pp. 503-512.

[16] E. C. Titchmarsh, *The Theory of Functions*, Oxford, 1932.

[17] ———, *Theory of Fourier Integrals*, Oxford, 1937.

[18] H. Whitney, "On ideals of differentiable functions," *American Journal of Mathematics*, vol. 70 (1948), pp. 635-658.

[19] N. Wiener, *The Fourier Integral and Certain of its Applications*, Cambridge, 1933.

ON UNIFORM DINI CONDITIONS IN THE THEORY OF LINEAR PARTIAL DIFFERENTIAL EQUATIONS OF ELLIPTIC TYPE.*

By PHILIP HARTMAN and AUREL WINTNER.

1. Introduction. Since Hölder's thesis [8] and the subsequent work of Korn and Lichtenstein (for references, cf. [14]) dealing with existence theorems for linear elliptic partial differential equations of second order, it has become standard practice to assume a uniform Hölder degree of continuity for the data. That is, it is assumed that the given functions (or suitable derivatives of them) satisfy a condition of the form

$$|f(P)-f(Q)| \leqq \text{const.}\, |PQ|^{\lambda}, \qquad 0<\lambda<1. \tag{1}$$

The proof of the existence theorems then depends on obtaining an a priori estimate for the degree of continuity of the second order derivatives of the solutions of the given, or of a related, partial differential equation. This a priori estimate is also of the form (1) (cf. [10], pp. 25-32); it allows the definition, and the proof for the convergence, of a sequence of successive approximations (cf. [11]).

It will be shown in the present paper that, in some of the existence theorems, the uniform Hölder degree of continuity for the data can be replaced by assumptions of a uniform degree of continuity,

$$|f(P)-f(Q)| \leqq \text{const.}\, \alpha(|PQ|), \tag{2}$$

where $\alpha=\alpha(r)$ is a monotone function satisfying Dini's condition (cf. [4], p. 102 and [5], pp. 210-219), that is,

$$\int_{+0} r^{-1}\alpha(r)\,dr<\infty, \qquad d\alpha(r)\geqq 0. \tag{3}$$

The proof of the resulting existence theorems will not depend on successive approximations. They will depend on an a priori estimate of the degree of continuity of the form

$$|f(P)-f(Q)| \leqq K\cdot \text{Const.}\, \beta(|PQ|) \tag{4}$$

for the second order partial derivatives of solutions. In (4), $\beta=\beta(r)$ is a function determined by $\alpha(r)$ and satisfying

$$\beta(r)\rightarrow 0 \text{ as } r\rightarrow 0, \tag{5}$$

---

* Received December 27, 1954.

Const. is a function of const. in (2), and K is a constant depending on the diameter of the domain under consideration. (In the case (1), $a(r)=\beta(r)=r^{\lambda}$, and K is a Korn constant.) The fact that $\beta(r)$ need not satisfy a Dini condition means that a sequence of successive approximations cannot in general be formed; cf. the example in Section 13 below. The existence proof can, however, be completed by using the method of equicontinuous functions.

Dini's results [5], involving a condition of the type (2), (3) in the theory of elliptic partial differential equations, antedates the work of Korn and Lichtenstein. But Dini only considers the simple equation $u_{xx}+u_{yy}+D_1u_x+D_2u_y+Du=E$. Dini's conditions are not expressed in the form (2), (3); instead of this, he merely requires that, uniformly in P,

$$(2') \qquad \infty > \int_{|PQ|<r} |f(P)-f(Q)| \cdot |P-Q|^{-n}\, dQ \to 0 \text{ as } r\to 0,$$

where $n=2$ is the number of independent variables and dQ is the n-dimensional element of volume (area). This milder condition (2′) can replace (2), (3) below if it is assumed, in addition, that, uniformly in P,

$$(3') \qquad |f(P)-f(Q)|\log|PQ| \to 0 \text{ as } Q\to P.$$

(For conditions, sharper than (2′), for the existence, but not for the continuity, of the second derivatives of a logarithmic potential, cf. [15].)

2. Statement of the results. In this paper, the method described in connection with (2)-(5), a method which has other applications also, will be used to give a detailed proof of the following theorem:

(*) *Let* $(g_{ik})=(g_{ik}(x,y))$ *be a 2 by 2, positive definite, symmetric matrix defined in a vicinity of* $(x,y)=(0,0)$ *in such a way that the matrix elements* $f=g_{11}, g_{12}=g_{21}, g_{22}$ *satisfy* (2), *where* $a=a(r)$ *is a continuous monotone function satisfying* (3). *Then, in a vicinity of* $(x,y)=(0,0)$, *there exist mappings* $u=u(x,y)$, $v=v(x,y)$ *of class* C^1 *satisfying*

$$(6) \qquad \partial(u,v)/\partial(x,y)\neq 0$$

and transforming the Riemann metric

$$(7) \qquad ds^2=g_{11}dx^2+2g_{12}dxdy+g_{22}dy^2$$

into the conformal normal form

$$(8) \qquad ds^2=\tau(du^2+dv^2), \qquad \tau=\tau(u,v)>0.$$

In addition, the first order partial derivatives of u and v will satisfy a condition of the type (4)-(5), where $\beta(r)$ is determined by $\alpha(r)$, the Const. in (4) is a function of the const. in (2) and of the bounds for g_{ik}, $(g_{11}g_{22}-g_{12}^2)^{-1}$ and g_{11}^{-1}, finally K is a constant (depending on $\alpha(r)$ and the diameter of the (x,y)-neighborhoods involved).

When $\alpha(r)=r^\lambda$, the assertion (*) reduces to a result of Lichtenstein ([13]; for another proof, cf. [2]). It is known ([7]; cf. [3], Appendix) that (*) is false if it is only assumed that the g_{ik} are continuous.

The partial differential equations connected with the problem of transforming (7) into (8) are of the form

$$v_x=-B_2u_x-Cu_y, \qquad v_y=Au_x+B_1u_y, \tag{9}$$

where

$$A=g_{22}/g, \quad B_1=B_2=-g_{12}/g, \quad C=g_{11}/g, \quad g=(\det g_{ik})^{\frac{1}{2}}. \tag{10}$$

Following the procedure of Lichtenstein [13] to obtain a solution u, v of (9), it will be shown that if R is sufficiently small, then the first boundary value problem of

$$(Au_x+B_1u_y)_x+(B_2u_x+Cu_y)_y=0 \tag{11}$$

on $x^2+y^2\leqq R^2$ has a C^1-solution, provided that the assigned boundary function is sufficiently smooth. By a C^1-solution $u=u(x,y)$ of (11) is meant a function of class C^1 on $x^2+y^2\leqq R^2$ which satisfies

$$\int_J (B_2u_x+Cu_y)\,dx-(Au_x+B_1u_y)\,dy=0 \tag{12}$$

for every piecewise smooth Jordan curve J in $x^2+y^2\leqq R$ (and which reduces to the given boundary values on $x^2+y^2=R^2$).

In the considerations below, it will not be assumed that $B_1=B_2$, as in (10), but only that the functions

$$a=A, \quad b=\tfrac{1}{2}(B_1+B_2), \quad c=C \tag{13}$$

satisfy

$$ac-b^2>0 \quad \text{and} \quad a>0. \tag{14}$$

In other words, it will be assumed that (9) is elliptic but not that it is self-adjoint; cf. [1]. It will be clear from the proof that the method can be applied to the inhomogeneous equation

$$(Au_x+B_1u_y)_x+(B_2u_x+Cu_y)_y+Du=E, \tag{15}$$

which in an integrated form is equivalent to

$$(16)\qquad \int_J (B_2u_x + Cu_y)\,dx - (Au_x + B_1u_y)\,dy = \int\int_T (Du - E)\,dxdy,$$

where T is the interior of the arbitrary piecewise smooth Jordan curve J.

(**) *Let* $f = A(x,y)$, $B_1(x,y)$, $B_2(x,y)$, $C(x,y)$ *be functions on* $x^2 + y^2 \leqq R^2$ *satisfying* (2)-(3), *let* $D(x,y)$, $E(x,y)$ *be continuous, and let* $\phi(x,y)$ *be of class* C^2 *there. Then there exists a number* R_0 (> 0), *depending on* const. *in* (2) *and on the bounds of* $|A|$, $|B_1|$, $|B_2|$, $|C|$, $|D|$, *with the property that if* $R \leqq R_0$, *then* (15) (*that is,* (16)) *has a* C^1-*solution* $u = u(x,y)$ *on* $x^2 + y^2 \leqq R^2$ *satisfying the boundary condition* $u = \phi$ *on* $x^2 + y^2 = R^2$.

(*This solution is of class* C^2 *if* A, B_1, B_2, C *are of class* C^1 *and their partial derivatives satisfy conditions of the type* (2)-(3) *and if* D, E *and the second order partial derivatives of* ϕ *satisfy conditions of the type* (2)-(3).)

The main interest in this theorem lies, of course, in the existence of non-constant solutions for (15), rather than in the solvability of a boundary value problem. For references to known results which depend on (1), rather than (2)-(3), cf. [14], pp. 1294-1927. The last (parenthetical) part of (**) will be seen (Section 14) to be a consequence of a modification of the proof of the first part; the conditions (2)-(3) on D, E and on the derivatives of A, B_1, B_2, C and ϕ can be relaxed to a uniform Dini condition (2′), where $n = 2$.

The assertions concerning (11) and (15) can be extended to the case of elliptic partial differential equations of second order in more than two independent variables. In this case, the definition of a C^1-solution of the analogue of (11) or (15) is not given by (12) or (16) but by an integro-differential equation involving Green's functions; cf. Section 7 below.

The proof for the existence of C^1-solutions of (11) will depend on variants of Korn's procedures [10] (adapted to (2), instead of (1)), as modified by E. Hopf [9]. While *loc. cit* [9], pp. 202-206, integrals depending only on the logarithmic singularity of the Green functions are dealt with, below it will be necessary to deal with the Green functions themselves, in order to obtain estimates which are valid inside (and on) the boundary curve of the domain.

Although the ideas underlying the proof of (*), as indicated in Section 1, are quite simple, the details will be involved. The arrangement of the proof of (*) will be as follows:

In Section 3, the main result (Lemma 1), dealing with a priori estimates for the degree of continuity of the partial derivatives of C^1-solutions of (11), will be stated, but not proved. Assuming Lemma 1, the existence theorem (Lemma 2) for the first boundary value problem belonging to (11) will be stated and proved in Section 4. A proof of (*) will then be deduced from this result in Section 5.

The proof of the main Lemma 1 will be given in Sections 6-12. In Section 6, the sheaf of Green functions will be introduced and their desired properties will be stated, but not proved, in Lemma 3. In Section 6, there will also be stated the analogues (Lemmas 4 and 5) of Hopf's variants of Korn's results. In Section 7, following Lichtenstein ([13], Section 2), the differential equation (11) will be transformed into an integro-differential equation involving the Green functions. Assuming Lemmas 3, 4, 5 of Section 6, the main Lemma 1 will be proved in Section 8. Finally, Lemma 3 will be proved in Sections 9-11 and Lemmas 4, 5 in Section 12.

Section 13 will contain an example mentioned in Section 1 in connection with successive approximations and (2)-(3).

3. A priori estimates. Let $\alpha = \alpha(r)$, where $0 \leqq r < \infty$, be a continuous monotone function satisfying (3). In terms of $\alpha(r)$ and a given $R > 0$, define $\beta(r) = \beta(r; R)$ as follows:

$$\beta(r) = (r + \alpha(r)) \log (8R/r) + \int_0^{4r} r^{-1}\alpha(r)dr + r\int_r^{2R} r^{-2}\alpha(r)dr \tag{17}$$

for $0 < r \leqq 2R$ (the choice of the factor 8 in $\log 8R/r$ assures that $\log 8R/r \geqq 1$ for $0 < r \leqq 2R$). By virtue of (3),

$$\beta(r) \to 0 \text{ as } r \to 0. \tag{18}$$

Let $f = A(x,y)$, $B_1(x,y)$, $B_2(x,y)$, $C(x,y)$ be real-valued functions defined on $\mathrm{C} + \mathrm{D}$, where

$$\mathrm{C}: x^2 + y^2 = R^2 \quad \text{and} \quad \mathrm{D}: x^2 + y^2 < R^2, \tag{19}$$

and let these four functions satisfy

$$|f(P) - f(Q)| \leqq k_1\alpha(|PQ|). \tag{20}$$

Let the constant k_1 be chosen so large that

$$|f(P)| \leqq k_1. \tag{21}$$

If $a = a(x,y)$, $b(x,y)$, $c(x,y)$ are defined by (13), suppose that (14) holds. Let k_1 be so large that

$$d^2 = ac - b^2 \geqq 1/k_1 \; (> 0), \qquad a \geqq 1/k_1 \; (> 0). \tag{22}$$

Let $\phi = \phi(x, y)$ be a (real-valued) function of class C^2 on C + D. Let the constant k_2 be chosen so that if f denotes ϕ or any of its first and second order partial derivatives, then

$$|f(P)| \leqq k_2. \tag{23}$$

The main lemma will be the following assertion, supplying an a priori estimate for the degree of continuity of first order derivatives of C^1-solutions of (11):

Lemma 1. *There exists a pair of functions of k_1, say $R_0(k_1)$ and $c_1(k_1)$, with the property that if $R \leqq R_0$ and if $u = u(x, y)$ is a C^1-solution of* (11) *on* C + D *satisfying the boundary condition*

$$u = \phi \text{ on } \mathrm{C}, \tag{24}$$

then the first order partial derivatives $f = u_x, u_y$ are subject to the estimates

$$|f| \leqq c_1(k_1) k_2, \tag{25$_1$}$$

$$|f(P) - f(Q)| \leqq c_1(k_1) k_2 \beta(|PQ|) \tag{25$_2$}$$

on C + D, *where $\beta(r) = \beta(r; R)$ is given by* (17).

Remark 1. Note that $c_1(k_1)$ does not depend on R ($\leqq R_0$).

Remark 2. It will be clear from the proof that conditions (2)-(3) on $f = A, B_1, B_2, C$ can be relaxed to (2′), (3′) in the following sense: Let E = E(r, P) denote the set of points Q in $x^2 + y^2 \leqq R^2$ satisfying $|Q - P| \leqq r$. Assume that

$$\epsilon_1(r) = \text{l. u. b.} \iint_{\mathrm{E}} |f(P) - f(Q)| \cdot |P - Q|^{-2} dQ \qquad (dQ = dxdy)$$

exists, where the l. u. b. refers to P in $x^2 + y^2 \leqq R^2$ and $f = A, B_1, B_2$ and C, and satisfies $\epsilon_1(r) \rightarrow 0$ as $r \rightarrow 0$. Then

$$\epsilon_2(r) = \text{l. u. b. } r \iint_{\mathrm{D-E}} |f(P) - f(Q)| \cdot |P - Q|^{-3} dQ$$

(exists and) satisfies $\epsilon_2(r) \rightarrow 0$ as $r \rightarrow 0$. Assume that

$$\epsilon_3(r) = \text{l. u. b. } (\log |PQ|^{-1}) |f(P) - f(Q)|,$$

(exists and) satisfies $\epsilon_2(r) \rightarrow 0$ as $r \rightarrow 0$, where l. u. b. refers to P in $x^2 + y^2 \leqq R^2$, Q in E(r, P) and $f = A, B_1, B_2, C$. Then Lemma 1 remains true if (17) is replaced by

$$\beta(r) = r \log (8R/r) + \epsilon_1(r) + \epsilon_2(r) + \epsilon_3(r).$$

Remark 3. It does not follow from the statement of Lemma 1 that if $\alpha(r)=r^{\lambda}$, where $0<\lambda<1$, then $\beta(r)$ can be chosen to $\beta(r)=r^{\lambda}$ (Korn). That this is so follows, however, from the fact that if (17) satisfies a Dini condition,

$$\int_{+0} r^{-1}\beta(r)\,dr<\infty,$$

then Lemma 1 can be improved by omitting the term $\alpha(r)\log 1/r$ in (17). The proof of this fact is similar to the application (p. 209) of the *Korollar* (p. 204) in [9].

It will be clear from the considerations to follow that it is sufficient to prove Lemma 1 for the case that A, B_1, B_2, C and ϕ are smooth; so that u is of class C^2.

4. Existence theorem for (11). The existence theorem, mentioned for (11) above, is the following Lemma 2. The assumptions on A, B_1, B_2, C and ϕ are those mentioned in connection with (19)-(22) and (23), respectively.

LEMMA 2. *If* $R\leqq R_0$, *there exists one and only one* C^1*-solution* $u=u(x,y)$ *of* (11) *on* C + D *satisfying the boundary condition* (24).

The existence assertion in this lemma will be deduced from the fact that Lemma 2 is known if A, B_1, B_2, C and ϕ are smooth (say, analytic); cf. [13], Sections 2-4 (where it is assumed that $B_1\equiv B_2$, but this assumption is not actually used).

There exist sequences $\{A_n\}$, $\{B_{1n}\}$, $\{B_{2n}\}$, $\{C_n\}$ of analytic functions on C + D satisfying, as $n\to\infty$,

$$(26_1)\qquad A_n\to A,\quad B_{1n}\to B_1,\quad B_{2n}\to B,\quad C_n\to C$$

uniformly on C + D and such that the inequalities (19)-(22), where k_1 is independent of n, are satisfied (the existence of such approximating sequences follows, for instance, from the theory of Fourier series). Similarly, there exists a sequence $\{\phi_n\}$ of smooth functions in C + D satisfying, as $n\to\infty$,

$$(26_2)\qquad \phi_n\to\phi$$

uniformly on C + D and such that ϕ_n and its first and second order partial derivatives satisfy (23), where k_2 is independent of n.

Let (11_n) denote the partial differential equation (11) in which A, B_1, B_2, C are respectively replaced by A_n, B_{1n}, B_{2n}, C_n and let (12_n) denote the integral relation corresponding to (12). Finally, let (24_n) denote the boundary condition in which ϕ is replaced by ϕ_n. If $u = u_n(x, y)$ denotes the solution of the boundary value problem (11_n), (24_n) and if $R \leqq R_0$, it follows from Lemma 1 and from (26_2) that the sequences $\{u_n\}$, $\{u_{nx}\}$, $\{u_{ny}\}$ are uniformly bounded and equicontinuous on C + D.

Consequently, after a selection and a renumbering of a subsequence, it can be supposed that, as $n \rightarrow \infty$,

$$(27) \qquad u = \lim u_n, \qquad u_x = \lim u_{nx}, \qquad u_y = \lim u_{ny}$$

exist uniformly on C + D. It follows from (26_2) and (24_n) that (24) holds. From (26_1), (27) and a term-by-term integration of (12_n), as $n \rightarrow \infty$, it is seen that (12) holds for every piecewise smooth Jordan curve J in D + C. This completes the existence proof.

The proof of uniqueness follows standard procedures for the proof of the weak maximum principle. The identity (12) in J implies the Green relation

$$\int_{\mathrm{J}} u\{(B_2u_x + Cu_y)dx - (Au_x + B_1u_y)dy\}$$
$$= -\iint_{\mathrm{T}} (au_x^2 + 2bu_xu_y + cu_y^2)dxdy,$$

where the double integral is over the interior T of J; cf. [6], Lemma, p. 761. Choose J = C and suppose that $u = 0$ on C. Then

$$\iint_{\mathrm{D}} (au_x^2 + 2bu_xu_y + cu_y^2)\,dxdy = 0;$$

so that, by (14), $u_x \equiv u_y \equiv 0$ on C + D, hence $u \equiv 0$. Thus (11) has exactly one C^1-solution satisfying the boundary condition $u = 0$ on C. In view of the linearity of (11), this completes the proof of the uniqueness in Lemma 2.

5. Proof of (*). Assertion (*) can now be proved by arguments of Lichtenstein [13], Section 5; its proof will be given only for the sake of completeness.

In terms of the g_{ik}, define A, $B_1 = B_2$, C by (10). Then, by assumption, $f = A_1$, $B_1 = B_2$, C satisfy (20)-(22) on C + D for suitable choices of R and k_1. Thus, if R is sufficiently small and ϕ is suitably chosen, the boundary value problem (11), (24) has a (unique) C^1-solution $u = u(x, y)$. It follows from the integrated form (12) of (11) that there exists a function $v = v(x, y)$ (unique up to an additive constant) which is of class C^1 on C + D and satisfies the system (9).

Hence, all that remains to be shown is that u and v satisfy (6) if R is sufficiently small and $\phi(x,y)$ is suitably chosen. To this end, choose a $k_2 > 0$, and let $\phi(x,y)$ be a function of class C^2 on C + D such that ϕ and its first and second order partial derivatives satisfy (23) and, in addition, the tangential derivative of ϕ at some point of C is not 0. For example, let the point (x,y) be $(0,R)$ and let $|\phi_x(0,R)| = k_2 \neq 0$. Since (24) implies that the tangential derivatives of u and ϕ coincide on C, it follows that $|u_x(0,R)| = k_2 \neq 0$. Since $c_1(k_1)$ in (25_2) does not depend on R $(\leqq R_0)$, it follows from the definition of $\beta(r) = \beta(r;R)$ that $u_x(x,y) \neq 0$ in C + D if R is sufficiently small.

The inequality $u_x(x,y) \neq 0$ on C + D implies (6), in view of (9) and (14). This proves (*).

6. Green's functions. In what follows, real- or complex-valued functions $f(x,y)$ of (x,y) will be denoted by $f(z)$. Correspondingly, $f_s(z)$, where $s = x$ or $s = y$, denotes $\partial f(x,y)/\partial s$.

Let a, b, c, d be real constants satisfying (14). Let $\zeta = \xi + i\eta$, $z = x + iy$ and let $G(\zeta, z) = G(\zeta, z; a, b, c)$ be the Green function of

$$aw_{xx} + 2bw_{xy} + cw_{yy} = 0 \tag{28}$$

on the circle D belonging to the boundary condition

$$w = 0 \text{ on C.} \tag{29}$$

Thus, if $h = h(x,y)$ is a sufficiently smooth function on C + D, then

$$w(x,y) = \frac{1}{2\pi} \iint_{D} h(\zeta)\, G(\zeta, z)\, d\xi d\eta \tag{30}$$

is the solution of

$$aw_{xx} + 2bw_{xy} + cw_{yy} = h \tag{31}$$

in D which vanishes on C.

The proof of Lemma 1 will be made to depend on some properties of the sheaf of Green functions $G(\zeta, z; a, b, c)$, where the constants (a, b, c) are subject to the inequalities

$$1/k_1 \leqq a, c \leqq k_1, \qquad d^2 = ac - b^2 \geqq 1/k_1, \tag{32}$$

in which k_1 is a positive constant.

For the purpose of denoting partial derivatives of G in Lemma 3, let $\sigma = \xi$ or $\sigma = \eta$, and $s, t = x$ or y, finally $e = a$, b or c.

LEMMA 3. *There exists a constant* $c_2 = c_2(k_1)$, *independent of* R, *such that*

$$(33)\qquad |G_s| \leqq c_2/|\zeta - z|, \qquad |G_{se}| \leqq c_2/|\zeta - z|,$$

$$(34)\qquad |G_{\sigma s}| \leqq c_2/|\zeta - z|^2, \qquad |G_{\sigma se}| \leqq c_2/|\zeta - z|^2,$$

$$(35)\qquad |G_{\sigma st}| \leqq c_2/|\zeta - z|^3,$$

for $|\zeta| < R$, $|z| < R$, $\zeta \neq z$.

Lemma 3 will be proved in Sections 9-11 below.

The following two lemmas are analogous to E. Hopf's variants ([9], pp. 202-206) of Korn's results ([10], pp. 25-32).

LEMMA 4. *Let* $H(\zeta, z; z_0)$ *be defined for* $\zeta\ (\neq z)$, z, z_0 *subject to* $|\zeta| < R$, $|z| < R$, $|z_0| < R$, *let* H *possess there continuous partial derivatives with respect to* x *and* y, *and let it satisfy, for some constant* k_3,

$$(36)\quad |H| \leqq k_3/|\zeta - z|, \quad |H(\zeta, z; z_0) - H(\zeta, z; z^0)| \leqq k_3\alpha(|z_0 - z^0|)/|\zeta - z|,$$

$$(37)\qquad |H_s| \leqq k_3/|\zeta - z|^2 \text{ for } s = x \text{ and } s = y.$$

Then, if $h(x, y)$ *is a continuous function on* C + D *and* k_4 *is a constant satisfying*

$$(38)\qquad |h| \leqq k_4,$$

the function

$$(39)\qquad \nu(z_0) = \iint_{\mathrm{D}} h(\zeta) H(\zeta, z_0; z_0)\, d\xi d\eta$$

satisfies

$$(40)\qquad |\nu(z_0) - \nu(z^0)| \leqq K k_3 k_4 \beta^1(|z_0 - z^0|),$$

where K *is an absolute constant and*

$$(41)\qquad \beta^1(r) = \beta^1(r; R) = R\alpha(r) + r \log(8R/r).$$

LEMMA 5. *Let* $L(\zeta, z; z_0)$ *be defined for* $\zeta\ (\neq z)$, z, z_0 *subject to* $|\zeta| < R$, $|z| < R$, $|z_0| < R$, *let* L *possess continuous partial derivatives with respect to* x *and* y *there and let it satisfy, for some constant* k_5,

$$(42)\quad |L| \leqq k_5/|\zeta - z|^2, \quad |L(\zeta, z; z_0) - L(\zeta, z; z^0)| \leqq k_5\alpha(|z_0 - z^0|)/|\zeta - z|^2,$$

$$(43)\qquad |L_s| \leqq k_5/|\zeta - z|^3 \text{ for } s = x \text{ and } s = y.$$

Let $h = h(x, y)$ *be continuous on* C + D *and let* k_4 *satisfy* (38). *Let* $l = l(x, y)$ *be defined on* C + D *and satisfy*

$$(44)\qquad |l(z_1) - l(z_2)| \leqq k_6\alpha(|z_1 - z_2|).$$

Then the function

$$(45)\qquad \mu(z_0)=\iint\limits_{D}(l(\zeta)-l(z_0))h(\zeta)L(\zeta,z_0;z_0)\,d\xi d\eta$$

satisfies

$$(46)\qquad |\mu(z_1)-\mu(z_2)|\leqq Kk_4k_5k_6\beta^2(|z_1-z_2|),$$

where K is an absolute constant and $\beta^2(r)=\beta^2(r;R)$ is defined by

$$a(r)\{\int_0^{2r}r^{-1}a(r)\,dr+\log(2R/r)\}+\int_0^{4r}r^{-1}a(r)\,dr+r\int_r^{2R}r^{-2}a(r)\,dr.$$

The proof of Lemma 4 is similar to, but simpler than, the proof of Lemma 5 and will therefore be omitted, Lemma 5 will be proved in Section 12 below.

Granting Lemmas 3, 4 and 5, the proof of Lemma 1 will be given in the next two sections.

7. The integro-differential equation. In this section and the next, it is assumed that A, B_1, B_2, C and ϕ are smooth (say, analytic). Let $u=u(x,y)$ be a solution of the boundary value problem (11), (24). Put

$$(47)\qquad w=u-\phi,$$

so that (24) gives

$$(48)\qquad w=0 \text{ on } C$$

and (11) becomes

$$(49)\qquad (A(w+\phi)_x+B_1(w+\phi)_y)_x+(B_2(w+\phi)_x+C(w+\phi)_y)_y=0.$$

Let $z_0=x_0+iy_0$ be a point of D. The differential equation (49) for w can be written as

$$(50)\qquad a_0w_{xx}+2b_0w_{xy}+c_0w_{yy}=h,$$

where

$$(51)\qquad \begin{aligned} h&=h_1-h_{2x}-h_{3y},\quad h_1=-a_0\phi_{xx}-2b_0\phi_{xy}-c_0\phi_{yy},\\ h_2&=(A-A_0)(w+\phi)_x+(B_1-B_{10})(w+\phi)_y,\\ h_3&=(B_2-B_{20})(w+\phi)_x+(C-C_0)(w+\phi)_y;\end{aligned}$$

(the superscript 0 denotes that the argument of the function is $z_0=x_0+iy_0$). In view of (48) and (50),

$$(52)\qquad 2\pi w(z)=\iint\limits_{D}h(\zeta)G(\zeta,z;a_0,b_0,c_0)\,d\xi d\eta.$$

Since G vanishes on the boundary of D, partial integrations of the terms involving h_{2x}, h_{3y} transform the last formula into

$$(52') \qquad 2\pi w(z) = \iint\limits_{D} (h_1 G + h_2 G_\xi + h_3 G_\eta)\, d\xi d\eta$$

(the fact that the singularity of G at $\zeta = z$ is of logarithmic order assures that this partial integration is permissible).

Since h_2, h_3 vanish at $z = z_0$ and $f = A, B_1, B_2, C$ satisfy (20), the Dini condition (3) and standard procedures in potential theory show that, if $s = x$ or $s = y$,

$$(53) \qquad 2\pi w_s(z_0) = \iint\limits_{D} (h_1 G_s + h_2 G_{\xi s} + h_3 G_{\eta s})\, d\xi d\eta,$$

where the argument of h_1, h_2, h_3 is ζ and that of the partial derivatives of G is $(\zeta, z_0; a_0, b_0, c_0)$.

8. Proof of Lemma 1. For a bounded function f on D, let

$$(54) \qquad \| f \| = \text{l. u. b.} \, | f | \text{ on D.}$$

The symbol c_1 will denote a constant depending on k_1 and independent of R $(\leqq R_0)$, c_2 a constant depending only on k_1, and K an absolute constant

From the definition of h_1 and from (21), (23) and (33),

$$\Big| \iint\limits_{D} h_1 G_s d\xi d\eta \Big| \leqq c_2 k_1 k_2 2\pi R.$$

From the definition of h_2 and from (21), (23) and (34),

$$\Big| \iint\limits_{D} h_2 G_{\xi s} d\xi d\eta \Big| \leqq c_2 k_1 (\| w_x \| + \| w_y \| + 2k_2)(2\pi) \int_0^{2R} r^{-1} \alpha(r)\, dr.$$

A similar estimate holds for the third term on the right of (53). Hence $\| w_x \| + \| w_y \|$ does not exceed

$$2c_2 k_1 k_2 (R + 4 \int_0^R r^{-1} \alpha(r)\, dr) + 2c_2 k_1 (\| w_x \| + \| w_y \|) \int_0^{2R} r^{-1} \alpha(r)\, dr.$$

Thus, if R is so small that

$$(55) \qquad 2c_2 k_1 \int_0^{2R} r^{-1} \alpha(r)\, dr \leqq \tfrac{1}{2}, \qquad c_2 = c_2(k_1),$$

it follows that

$$(56)\qquad \| w_x \| + \| w_y \| \leqq 4c_2k_1k_2(R + 4\int_0^R r^{-1}\alpha(r)\,dr).$$

If $R = R_0(k_1)$ is a fixed positive number satisfying (55), then the assertion of Lemma 1 concerning (25_1) follows from (47) and (56).

Lemma 4 is applicable to the first term on the right of (53), that is, to

$$(57)\qquad \nu(z_0) = \iint_D h_1(\zeta)\, G_s(\zeta, z_0; a_0, b_0, c_0)\, d\xi d\eta,$$

if G_s, h_1 are identified with H, h, respectively. In fact, the first inequality of (36) is a consequence of the first in (33), with $k_3 = c_2$; the second in (36) follows, with $k_3 = 3c_2k_1$, from the second in (33), and from (13) and (20). The second line in (51) shows that k_4 in (38) can be replaced by $2k_1k_2$. Thus the assertion of Lemma 4 implies that

$$(58)\qquad | \nu(z_0) - \nu(z^0) | \leqq Kc_1k_2\beta^1(| z_0 - z^0 |) \leqq Kc_1k_2\beta(| z_0 - z^0 |).$$

If h_2, h_3 in (53) are replaced by their equivalent expressions given by (51), the last two terms of (53) become four terms, to each of which Lemma 5 is applicable. For the sake of definiteness, one of the four terms, say the term

$$(59)\qquad \mu(z_0) = \iint_D (A - A_0)(w + \phi)_x G_{\xi s}(\zeta, z_0; a_0, b_0, c_0)\, d\xi d\eta,$$

will be considered. If L in Lemma 5 is identified with $G_{\xi s}$, it follows from (20), (34) and (35) that (42), (43) hold, with k_5 a constant (depending only on k_1). Let l be identified with A, so that (44) holds with $k_6 = k_1$. Finally, let h in (45) be identified with $(w + \phi)_x$ in (59), so that (38) holds with $k_4 = k_2 + \| w_x \|$. Thus Lemma 5 implies that

$$| \mu(z_0) - \mu(z^0) | \leqq Kc_1(k_2 + \| w_x \|)\beta(| z_0 - z^0 |),$$

where $c_1 = c_1(k_1) = c_1(k_1, R_0)$ and $\beta^2(r) \leqq c_1\beta(r)$ for a suitable choice of $c_1 = c_1(k_1)$. It follows from (56) that this can be written as

$$| \mu(z_0) - \mu(z^0) | \leqq Kc_1k_2(R + 4\int_0^R r^{-1}\alpha(r)\,dr)\beta(| z_0 - z^0 |).$$

Since the other terms resulting from (53) can be treated similarly, it is seen that

$$(60)\qquad | w_s(z_0) - w_s(z^0) | \leqq c_1k_2\beta(| z_0 - z^0 |),$$

if $c_1 = c_1(k_1) = c_1(k_1, R_0)$ and $R \leqq R_0$. The inequality (23), when applied to the case in which f is a second order partial derivative of ϕ, shows that

$$| \phi_s(z_0) - \phi_s(z^0) | \leqq k_2 | z_0 - z^0 |.$$

Since $\beta(r) \geqq r$ when $r \leqq 2R$, the assertion of Lemma 1 concerning (25_2) follows from (47) and (60).

This completes the proof of Lemma 1.

9. The inequalities (33_j), (34_j), (35_j), $j = 1, 2$. Lemma 3, that is, the properties (33)-(35) of G, will be deduced from properties of the Green function

$$\gamma(\mathbf{Z}, Z) = \log | (R^2 - \bar{Z}\mathbf{Z}) / R(Z - \mathbf{Z}) | \tag{61}$$

belonging, in $|Z| < R$, to Laplace's equation and the boundary condition $u = 0$ on $|Z| = R$.

The change of independent variables

$$U = x/a^{\frac{1}{2}}, \qquad V = (-bx + ay)/(a^{\frac{1}{2}}d) \tag{62}$$

transforms (28) into

$$w_{UU} + w_{VV} = h, \tag{63}$$

C into an ellipse C_0 and D into the interior D_0 of C_0. Let $g(\Omega, W)$ be the Green function belonging to (63) on D_0. Then

$$G(\zeta, z) = d^{-1} g(W(\zeta), W(z)), \text{ where } W = U + iV, \tag{64}$$

$W = W(z)$ is the affine transformation (62), and $d^{-1} = \partial(U, V)/\partial(x, y)$. Thus if $Z = Z(W) = Z(W; a, b, c)$ is a conformal mapping of D_0 onto the circle $|W| < R$, it follows that

$$g(\Omega, W) = \gamma(Z(\Omega), Z(W)). \tag{65}$$

The formulae (64), (65) give G explicitly in terms of (61) and a mapping function $Z = Z(W)$. The latter is expressible in terms of elliptic functions (cf. [16]) and can be chosen so as to be analytic in the real parameters a, b, c, as well as in W. In view of (64) and (65), a formula for G is

$$G(\zeta, z; a, b, c) = d^{-1} \gamma(\mathbf{Z}, Z), \tag{66}$$

where

$$Z = Z(W(z)), \qquad \mathbf{Z} = Z(W(\zeta)), \tag{67}$$

and γ is defined by (61).

The following notation is used below for the variables which occur:

$$z = x + iy, \quad W = U + iV, \quad Z = X + iY, \quad \zeta = \xi + i\eta, \quad \mathbf{Z} = \Xi + i\mathrm{H}.$$

The symbol c_2 will denote a positive constant, not always the same, depending only on k_1 and independent of R. The letter K will signify some positive (absolute) constant.

In order to follow the dependence of G on R, let $Z = Z^1(W)$ be the function $Z(W)$ for the case $R = 1$. Then $Z(W)$, belonging to R, is $Z(W) = RZ^1(W/R)$. Note that the affine transformation (62) does not depend on R and is linear and homogeneous, so that $W(z)/R = W(z/R)$. It follows that

$$(68_1)\ c_2^{-1} \leqq |Z_W| \leqq c_2; \quad (68_2)\ |Z_{WW}| \leqq c_2/R; \quad (68_3)\ |Z_{WWW}| \leqq c_2/R^2.$$

If $e = a$, b or c and $Z = Z(W, a, b, c)$, then

$$(69_1)\ |Z_e| \leqq c_2 R; \qquad (69_2)\ |Z_{We}| \leqq c_2; \qquad (69_3)\ |Z_{WWe}| \leqq c_2/R.$$

In particular, the ratio $|Z(W_1) - Z(W_2)|/|W_1 - W_2|$ is bounded from above and from below by c_2, c_2^{-1}, respectively. Hence (62) and (22) imply that

$$(70_1) \qquad c_2^{-1} \leqq |Z(W(z)) - Z(W(\zeta))|/|z - \zeta| \leqq c_2,$$

$$(70_2) \qquad |Z_e(W(z)) - Z_e(W(\zeta))| \leqq c_2 |z - \zeta|.$$

The inequalities (69) and (70_2), involving partial derivatives with respect $e = a$, b or c, were derived by considering $Z = Z(W, a, b, c)$ as a function of W and a, b, c. The function $Z(W(z)) = Z(W(z; a, b, c), a, b, c)$, where $W = W(z; a, b, c)$ is the affine transformation (62), depends however on a, b, c in a (slightly) more complicated fashion. The derivative of $Z(W(z))$ with respect to $e = a$, b or c is $Z_W W_e + Z_e$. Since W_e depends linearly on x and y, it follows that $|W_e| \leqq c_2 R$ and $|W_e(z) - W_e(\zeta)| \leqq c_2 |z - \zeta|$. Hence (69_1) and (70_2) hold if Z_e is also an abbreviation for $Z_W W_e + Z_e$. It is in this sense that Z_e will be used below. Similar remarks hold for (69_2) and (69_3).

Let the function (61) be written as $\gamma = \gamma_1 + \gamma_2 - \log R$, where

$$(71_1) \qquad \gamma_1 = -\log |\mathbf{Z} - Z|, \qquad\qquad (71_2) \qquad \gamma_2 = \log |R^2 - \bar{Z}\mathbf{Z}|.$$

Let the equations (66)-(67), where $\gamma = \gamma_1$ or $\gamma = \gamma_2$, define G_1 and G_2; so that $G = d^{-1}(G_1 + G_2 - \log R)$. The proofs of (33)-(35) will be given separately for G_1, G_2. The inequalities analogous to (33), (34), (35) for G_j will be referred to as (33_j), (34_j), (35_j), where $j = 1, 2$.

10. Proof of (33₁), (34₁), (35₁). If $s=x$ or $s=y$, then (66)-(67) give

$$G_{js}=\gamma_{jX}X_s+\gamma_{jY}Y_s, \tag{72}$$

where, for $j=1$,

$$\gamma_{1X}=(\Xi-X)/|\mathbf{Z}-Z|^2, \qquad \gamma_{1Y}=(\mathrm{H}-Y)/|\mathbf{Z}-Z|^2. \tag{73}$$

Since (68₁) implies that $|X_s|$, $|Y_s|\leqq c_2$, the first inequality in (33₁) follows from (67) and (70₁). The second inequality in (33₁) follows from (69₂) and (70₂).

In order to prove (34₁), differentiate (72) with respect to $\sigma=\xi$ or $\sigma=\eta$. This gives

$$G_{js\sigma}=\gamma_{jX\Xi}X_s(W(z))X_t(W(\zeta))+\gamma_{jX\mathrm{H}}X_x(W(z))Y_t(W(\zeta))+\cdots, \tag{74}$$

where $t=x$ or $t=y$ according as $\sigma=\xi$ or $\sigma=\eta$, and $+\cdots$ indicates two more terms. When $j=1$, the first factor of the first term on the right is

$$\gamma_{1X\Xi}=1/|\mathbf{Z}-Z|^2-2(\Xi-X)^2/|\mathbf{Z}-Z|^4; \tag{75}$$

the other factors $\gamma_{1X\mathrm{H}}$, $\gamma_{1Y\Xi}$, $\gamma_{1Y\mathrm{H}}$ are similar. Thus the first inequality in (34₁) follows from (67), (68₁) and (70₁). If $j=1$ and if (74) is differentiated with respect to $e=a$, b, or c, then (70₂) shows that $\gamma_{1X\Xi e},\cdots$ have majorants of the form $c_2/|z-\zeta|^2$, while (69₁) and (69₂) show that $X_s, X_{se},\cdots$ have majorants of the form c_2. This proves (34₁).

In order to prove (35₁), let $j=1$ and let (74) be differentiated with respect to $t=x$ or $t=y$. It is clear that the resulting terms involving a third order partial derivative of γ (with respect to X, Y, Ξ, H) are majorized by $c_2/z-\zeta|^3$, while, by (68₂), those involving a second order partial derivative of γ, and a second order partial derivative of X or Y, are majorized by $c_2/|z-\zeta|^2R$. Since $2R>|z-\zeta|$, the inequality (35₁) follows.

11. Proof of (33₂), (34₂), (35₂). The proofs of these inequalities will be similar to those of (33₁), (34₁), (35₂) but they will depend, in addition, to (68), (69) and (70), on two elementary inequalities and on some remarks on the inequalities (70) for the case when $|z|=R$ or $|\zeta|=R$. Both of the required elementary inequalities, which can be interpreted in terms of the geometry of the circle, will be proved for the sake of completeness.

The first elementary inequality is

$$R^{-1}|R^2-\bar{Z}\mathbf{Z}|\geqq|\mathbf{Z}-Z|/4, \qquad |Z|<R, |\mathbf{Z}|<R. \tag{76}$$

In order to verify (76), note that if Z_* is the mid-point of the line segment joining Z and $R^2/\bar{Z}$, then $|Z_*| > R$. Hence, the perpendicular bisector of this segment does not intersect the circle $|Z| = R$. Consequently,

$$|\mathbf{Z} - R^2/\bar{Z}| \geqq |\mathbf{Z} - Z|.$$

But $R^{-1}|R^2 - \bar{Z}\mathbf{Z}| = |Z/R| \cdot |\mathbf{Z} - R^2/\bar{Z}|$, and so (76) holds if $|Z| \geqq \frac{1}{2}R$. On the other hand, if $|Z| < \frac{1}{2}R$, then $R^{-1}|R^2 - \bar{Z}\mathbf{Z}| \geqq \frac{1}{2}R \geqq |\mathbf{Z} - Z|/4$. This completes the verification of (76).

The second elementary inequality is

$$R(R - |\mathbf{Z}|)/|R^2 - \bar{Z}\mathbf{Z}| \leqq 2. \tag{77}$$

The proof of this is obvious if it is remarked that the left side does not exceed $|R/Z|$.

The derivations of the inequalities (68)-(70) show that they are valid if the arguments z, ζ, W of the functions involved are on the boundaries C, C, C_0 of the appropriate circular or elliptical domain. This fact will now be used to derive the following consequences of (70_1), (70_2), respectively:

$$c_2^{-1} \leqq (R - |Z(W(z))|)/(R - |z|) \leqq c_2, \tag{78_1}$$

$$|\partial(|Z(W(z))|^2 - R^2)/\partial e| \leqq c_2(R - |Z|) \qquad (e = a, b, \text{or } c). \tag{78_2}$$

In order to verify the second inequality in (78_1), let z be fixed ($|z| < R$) and let z_0 be a point of C nearest z. Thus $|z - z_0| = R - |z|$. Then $|Z(W(z_0))| = R$, and so

$$|Z(W(z)) - Z(W(z_0))| \geqq R - |Z(W(z))|.$$

The second inequality in (70_1) shows that the expression on the left is majorized by $c_2|z - z_0| = c_2(R - |z|)$. Hence the second inequality in (78_1) follows. The first inequality in (78_1) is similarly proved.

If R^2 is replaced by $|Z(W(z_0))|^2$, then (78_2) follows from the case $\zeta = z_0$ of (70_2) (and the fact that $||f|_e| \leqq |f_e|$).

The inequalities (77) and (78_1) will be used together, in the form

$$R(R - |\zeta|)/|R^2 - \bar{Z}\mathbf{Z}| \leqq c_1, \tag{79}$$

where Z and $\mathbf{Z}$ are given by (67).

Finally,

$$R^2 - \bar{Z}\mathbf{Z} = \mathfrak{R} + i\mathfrak{I}, \tag{80}$$

where $\mathfrak{R}$ and $\mathfrak{I}$ are real and are given by

$$\begin{aligned} \mathfrak{R} &= (R^2 - |\mathbf{Z}|^2) + \Xi(\Xi - X) + \mathrm{H}(\mathrm{H} - Y), \\ \mathfrak{I} &= \mathrm{H}(X - \Xi) + \Xi(\mathrm{H} - Y) \end{aligned} \tag{81}$$

or, equivalently, by

$$\mathfrak{R} = (R^2 - |Z|^2) + X(X - \Xi) + Y(Y - \mathrm{H}), \ \mathfrak{J} = X(\mathrm{H} - Y) + Y(X - \Xi).$$

The proof of (33_2) can now proceed as follows:

In the case $j = 2$ of (72),

$$(82) \quad \gamma_{2X} = (-\Xi\mathfrak{R} + \mathrm{H}\mathfrak{J})/|R^2 - \bar{Z}\mathbf{Z}|^2; \quad \gamma_{2Y} = (-\mathrm{H}\mathfrak{R} - \Xi\mathfrak{J})/|R^2 - \bar{Z}\mathbf{Z}|^2.$$

Thus $|\gamma_{2S}| \leqq R/R^2 - \bar{Z}\mathbf{Z}|$, where $S = X$ or $S = Y$. Hence the first inequality in (33_2) follows from (76) and (70_1).

In order to prove the second part in (33_2), note that a differentiation of (72), where $j = 2$, with respect to e ($= a$, b or c) leads to terms each of which is majorized by one of the expressions

$$c_2R/|R^2 - \bar{Z}\mathbf{Z}|, \quad c_2R^2(R - |\mathbf{Z}|)/|R^2 - \bar{Z}\mathbf{Z}|^2, \quad c_2R^2|z - \zeta|/|R^2 - \bar{Z}\mathbf{Z}|^2,$$

by virtue of (81), (82) and (69), (70), (78). Hence the second inequality in (33_2) follows from (76) and (77).

This proof of (33_2) shows that (34_2), (35_2) can be proved in a manner similar to that used above for (34_1), (35_1), with $\mathfrak{R}$, $\mathfrak{J}$ in (81) playing the rôles of the real and imaginary parts, $\Xi - X$ and $\mathrm{H} - Y$, of $\mathbf{Z} - Z$. The details of these arguments will therefore be omitted.

Since the factor d^{-1} in (66) is simply the function $(ac - b^2)^{-\frac{1}{2}}$ of a, b, c, it is clear that Lemma 3 follows from the inequalities (33_j), (34_j), (35_j) for $j = 1, 2$.

12. Proof of Lemma 5. The function (45) can be written as $\mu(z_0, z_0)$, where

$$(83) \qquad \mu(z, z_0) = \iint\limits_{\mathrm{D}} (l(\zeta) - l(z))h(\zeta)L(\zeta, z; z_0)\,d\xi d\eta.$$

In order to prove Lemma 5, it is sufficient to verify inequalities of the type (46) for both

$$(84) \qquad |\mu(z, z_0) - \mu(z, z^0)|$$

and

$$(85) \qquad |\mu(z_2, z_0) - \mu(z_1, z_0)|.$$

The second inequality in (42) and the inequalities (38) and (44) show that (84) is majorized by

$$k_4k_5k_6a(|z_0 - z^0|) \iint\limits_{\mathrm{D}} a(|\zeta - z|)\,|\zeta - z|^{-2}\,d\xi d\eta.$$

The last integral does not exceed $2\pi \int_0^{2R} \alpha(r) r^{-1} dr$. Hence

$$(86) \qquad |\mu(z, z_0) - \mu(z, z^0)| \leqq 2\pi k_4 k_5 k_6 \alpha(|z_0 - z^0|) \int_0^{2R} \alpha(r) r^{-1} dr.$$

In order to appraise (85), let

$$(87) \qquad r = 2|z_1 - z_2|$$

and let $E = E(r)$ denote the portion of the circular disc $|z - z_1| < r$ contained in D,

$$(88) \qquad E\colon \ |z - z_1| < r, |z| < R.$$

Then the difference whose absolute value occurs in (85) can be written as

$$(89) \qquad \mu(z_2, z_0) - \mu(z_1, z_0) = I_1 + I_2 - (\Delta l) I_3,$$

with

$$(90) \qquad \begin{aligned} I_1 &= \Delta \iint_E J(\zeta, z; z_0)\, d\xi d\eta, \\ I_2 &= \iint_{D-E} (l(\zeta) - l(z_1)) h(\zeta) \Delta L(\zeta, z; z_0)\, d\xi d\eta, \\ I_3 &= \iint_{D-E} h(\zeta) L(\zeta, z_2; z_0)\, d\xi d\eta, \end{aligned}$$

where $J(\zeta, z; z_0)$ is the integrand in (83) and, for any function $g(z, \cdot \cdot \cdot)$ depending on z, the symbol Δg denotes the difference $g(z_2, \cdot \cdot \cdot) - g(z_1, \cdot \cdot \cdot)$.

By (38), (44) and the first inequality in (42),

$$\iint_E |J(\zeta, z; z_0)|\, d\xi d\eta \leqq k_4 k_5 k_6 \iint_E \alpha(|\zeta - z|) |\zeta - z|^{-2}\, d\xi d\eta.$$

If ζ is on E and if $z = z_1$ or $z = z_2$, then $|z - \zeta| \leqq 2r$. Hence

$$(91) \qquad |I_1| \leqq 2(k_4 k_5 k_6)(2\pi) \int_0^{2r} r^{-1} \alpha(r)\, dr.$$

It follows from (43) that $|\Delta L| \leqq \frac{1}{2} k_5 r / |z_* - \zeta|^3$, where z_* is some point on the line segment joining z_1 and z_2. If ζ is not in E, then $|z_* - \zeta| \geqq |z_1 - \zeta| - |z_1 - z_*|$. Since $|z_1 - z_*| \leqq \frac{1}{2} r \leqq \frac{1}{2} |z_1 - \zeta|$, it follows that $|z_* - \zeta| \geqq \frac{1}{2} |z_1 - \zeta|$. Hence $|\Delta L| \leqq 4 k_5 r / |z_1 - \zeta|^3$. Thus, by (38) and (44),

$$|I_2| \leqq 4(k_4k_5k_6)(2\pi)r\int_r^{2R} r^{-2}a(r)\,dr. \tag{92}$$

Finally, (38) and the first inequality in (42) imply that

$$|I_3| \leqq k_4k_5 \iint_{\mathrm{D-E}} |\zeta - z_2|^{-1}d\xi d\eta \leqq (k_4k_5)(2\pi)\int_{\frac{1}{2}r}^{2R} r^{-1}dr;$$

so that, by (44),

$$|(\Delta l)I_3| \leqq (k_4k_5k_6)(2\pi)a(\tfrac{1}{2}r)\log(4R/r). \tag{93}$$

By (89), (91), (92) and (93), the difference (85) is majorized by

$$K(k_4k_5k_6)\{a(\tfrac{1}{2}r)\log(4R/r) + \int_0^{2r} r^{-1}a(r)\,dr + r\int_{\frac{1}{2}r}^{2R} r^{-2}a(r)\,dr\},$$

where $\frac{1}{2}r = |z_1 - z_2|$. Hence, Lemma 5 follows from (86), by virtue of the definition of $\beta^2(r)$ in the formula line following (46).

13. An example. Let $F = F(r)$ be a continuous function on $0 \leqq r \leqq R$ satisfying

$$F(r) \geqq 0 \text{ according as } r \geqq 0 \tag{94}$$

and let S be the surface of revolution defined by

$$S\colon\ z = z(x,y) = \int_0^r F(r)\,dr \qquad (r = (x^2+y^2)^{\frac{1}{2}} \geqq 0). \tag{95}$$

(In this section, z is real and (x,y,z) are coordinates in space.) Thus S is a surface of class C^1. The element of arc-length on S is given by (7), where

$$g_{11} = 1 + F^2x^2/(x^2+y^2), \qquad g_{12} = F^2xy/(x^2+y^2),$$
$$g_{22} = 1 + F^2y^2/(x^2+y^2). \tag{96}$$

Let $a = a(r)$ be defined, for $0 \leqq r \leqq \frac{1}{2}$, by

$$a(r) = 1/\log^2 r \text{ or } a(r) = 0 \text{ according as } r \geqq 0 \tag{97}$$

and let

$$F(r) = -1/\log r \text{ or } F(r) = 0 \text{ according as } r \geqq 0. \tag{98}$$

Then (97) satisfies (3) and $f = F^2$ satisfies (2) on $x^2 + y^2 \leqq R^2$ $(\leqq 1/4)$.

It is easily verified that $f = g_{ik}$ also satisfy (2)-(3) for $i, k = 1, 2$, where g_{ik} is given by (96) and $a(r)$ by (97).

A mapping $u = u(x, y)$, $v = v(x, y)$ of class C^1 transforming (7) into the conformal normal form is given by

$$(99) \qquad u = \rho(r)x, \qquad v = \rho(r)y, \qquad r = (x^2 + y^2)^{\frac{1}{2}},$$

where, for $r > 0$,

$$(100) \qquad \rho(r) = r^{-1} \exp(-\int_r^R r^{-1}(1 + F^2(r))^{\frac{1}{2}} dr);$$

the conformal normal form (8) for ds^2 is

$$(101) \qquad ds^2 = \rho^{-2}(r)(du^2 + dv^2);$$

cf. [3], Appendix. The definition of $\rho(0)$ follows by continuity from (102) below; the fact that u, v are of class C^1 can be seen from the derivation of (102); finally, $\partial(u, v)/\partial(x, y) \neq 0$ is a consequence of $\rho > 0$.

Note that (100) satisfies $R\rho(r) = \exp \int_r^R \{r^{-1}(1 - (1 + F^2(r))^{\frac{1}{2}}\} dr$; so that

$$R\rho(r) = \exp(\int_0^R - \int_0^r) = \text{Const.} \exp - \int_0^r, \text{ where Const.} = \exp \int_0^R.$$

On writing $(1 + F^2)^{\frac{1}{2}}$ as $1 + \frac{1}{2}F^2(r) + \cdots$, it is seen that

$$(\text{Const.}/R)\rho^{-1}(r) = \exp \int_0^r (\tfrac{1}{2}r^{-1}F^2(r) + \cdots) dr.$$

Since $\int_0^r r^{-1}F^2(r) dr = -1/\log r$, by (98), it follows that

$$(102) \qquad \rho^{-1}(r) = (R/\text{Const.})\{1 - \tfrac{1}{2}(1 + o(1))/\log r)\}.$$

Hence, the factor $\rho^{-2}(r)$ in (101) satisfies, for small $r > 0$,

$$(103) \qquad \rho^{-2}(r) - \rho^{-2}(0) \geqq (\tfrac{1}{2}R/\text{Const.})^2(-1/\log r).$$

By the transformation rule for a metric tensor, there exists a constant $c > 0$ such that at least one of the four partial derivatives $f = u_x, u_y, v_x, v_y$ satisfies

$$(104) \qquad |f(x, y) - f(0, 0)| \geqq c/|\log r|.$$

On the other hand, the $\beta(r)$ in (17) belonging to (97) satisfies $\beta(r) \leqq C/|\log r|$ for a suitable choice of $C = C(R)$. Hence, by Lemma 1, $f = u_x, u_y, v_x, v_y$ satisfy

$$|f(P) - f(Q)| \leqq C/|\log|PQ||. \tag{105}$$

This example shows that, in general, the function $\beta(r)$ in Lemma 1 cannot be replaced by a "smaller" function. It also shows that, while $\alpha(r)$ satisfies the Dini condition (3), the function $\beta(r)$ need not, and that $\beta(r)$ cannot be replaced by a function which satisfies a Dini condition

$$\int_{+0} r^{-1}\beta(r)\,dr < \infty.$$

As observed in Section 1, it follows that when (1) is relaxed to (2)-(3), the usual method of successive approximations cannot supply existence theorems.

The example of a ds^2 given by (96) is a modification of one given by Lavrentieff [12], p. 420, and used in [3], Appendix, for a related purpose.

[In the last formula line of [3], p. 308, the term $x/2R$ should be $x^2/2R$.]

14. Remarks on the proof of (**). It is clear from the proof of (*) that the terms Du, E, those which occur in (15), but not in (11), cause no complications and that the proof of the main (non-parenthetical) part of (**) follows from that of Lemma 2. In fact, the same is true even if Du in (15) is replaced by $F_1u_x + F_2u_y + Du$, when D, F_1, F_2 are continuous. In this case, however, R_0 will depend also on the bounds for $|F_1|$, $|F_2|$.

Since analogous remarks apply to the last (parenthetical) part of (**), the proof of this part of (**) will be indicated only for the case (11) of (15), where $D = E = 0$. It will be shown that, under the conditions stated, the proof of Lemma 1 can be modified so as to lead to an a priori estimate for the degree of continuity, and to a priori bounds, for the second order partial derivatives of C^2-solutions $u = u(x, y)$ of (11).

To this end, consider the integro-differential equation (52) for w $(= u - \phi)$. Differentiation of (52) with respect to $s = x$ or $s = y$ leads to

$$2\pi w_s(z) = \iint_D h(\zeta)\, G_s(\zeta, z; a_0, b_0, c_0)\, d\xi d\eta. \tag{106}$$

This step can be made, since it is assumed that A, B_1, B_2, C are of class C^1 and that u, ϕ are of class C^2.

Since A, B_1, B_2, C are of class C^1, it follows from (*) that w_x, w_y satisfy a uniform Hölder condition of any order $\lambda < 1$ (with a Hölder constant depending only on λ, R_0 and the bounds for the partial derivatives of A, B_1, B_2, C). Hence, the motivation for (53) assures the validity of the formal differentiation of (106) with respect to $t=x$ or $t=y$ at the point $z=z_0$,

$$(107)\qquad 2\pi w_{st}(z_0)=h(z_0)\Big(\iint\limits_{D} G_s\,d\xi d\eta\Big)_t+\iint\limits_{D}(h(\zeta)-h(z_0))G_{st}\,d\xi d\eta,$$

where the argument of the partial derivative of G in the last integral is $(\zeta, z_0; a_0, b_0, c_0)$. In fact, this motivation is more readily seen if the terms h_{2x}, h_{2y} in h are written as

$$h_{2x}=A_x(w+\phi)_x+(A-A_0)(w+\phi)_{xx}+\cdots,$$
$$h_{3y}=B_{2y}(w+\phi)_x+(B_2-B_{20})(w+\phi)_{xy}+\cdots.$$

Since Lemma 3 holds if $\sigma=x$ or $\sigma=y$ (as well as $\sigma=\xi$ or $\sigma=\eta$ in (34)-(35)), the last formula and (107) show that, if R is sufficiently small, then the arguments leading to Lemma 1 supply a priori estimates for the degree of continuity and a priori bounds for w_{st}. The proof of Lemma 2 shows that under the assumptions of the last part of (**), C^1-solutions of (11) are of class C^2.

APPENDIX.

Potential Theory and Lebesgue Constants.

Let $f=f(x,y)$ be uniformly continuous on an open bounded domain D of the (x,y)-plane. Then a *formal* solution $\phi=\phi(x,y)$ of Poisson's equations

$$(1)\qquad \phi_{xx}+\phi_{yy}=f$$

on D is the logarithmic potential

$$(2)\qquad \phi(x,y)=\frac{1}{2\pi}\iint\limits_{D} f(u,v)\log r\,dudv,$$

where $r^2=(x-u)^2+(y-v)^2$. Actually, the second derivatives of the function (2) need not exist on D if $f(x,y)$, instead of satisfying something like a Hölder condition, is just uniformly continuous. In fact, Petrini [15], pp. 132-133, has shown that the derivative ϕ_{xx} of (22) will exist at a point (a,b) of D if and only if the limit

$$(3)\qquad \lim_{\epsilon\to 0}\iint\limits_{D-E(\epsilon;a,b)} f(u,v)(\log r)_{uu}\,dudv \text{ exists,}$$

where $E(\epsilon; a, b)$ denotes the interior of the circle of radius ϵ about (a, b), and that a corresponding criterion holds for ϕ_{yy} and ϕ_{xy} (or ϕ_{yx}), with $(\log r)_{uu}$ replaced by $(\log r)_{vv}$, $(\log r)_{uv} = (\log r)_{vu}$ in (3). (Incidentally, this implies that $\phi_{xy}(a, b)$ exists if and only if $\phi_{yx}(a, b)$ does [and is $\phi_{yx}(a, b)$]; in addition [cf. *loc. cit.*, p. 134], the existence of $\phi_{xx}(a, b)$ is equivalent to that of $\phi_{yy}(a, b)$).

Petrini observes ([15], p. 138) that if D contains $(0, 0)$ and if

$$(4) \qquad f(x, y) = x^2/(r^2 \log r) \qquad (\text{if } (x, y) \neq (0, 0), \text{ and } f(0, 0) = 0),$$

then (3) fails to hold at $(a, b) = (0, 0)$. This proves that the second derivatives of (2) need not exist when f is just continuous.

Another turn to Petrini's criterion (3) was given in [17]. This turn, which in [17] made possible the passage from (1) to the homogeneous equation

$$(5) \qquad \psi_{xx} + \psi_{yy} + f\psi = 0$$

(and which was further exploited in [7]), consists in the following observation: Instead of finding an explicit example (such as (4)) which violates Petrini's condition (3), it is sufficient to think of the limit (3) (if any) as a "singular integral" in the sense of Lebesgue, with

$$(6) \qquad \iint\limits_{D-E(\epsilon; a, b)} |(\log r)_{uu}|\, du dv$$

as the "ϵ-th Lebesgue constant." Since (6) tends to ∞ as $\epsilon \rightarrow 0$, it is sufficient to appeal to Lebesgue's classical "norm construction" (cf. his *Leçons sur les séries trigonométriques* (1906), pp. 85-87) in order to obtain a uniformly continuous $f(x, y)$ for which the logarithmic potential (2) fails to possess second derivatives.

The object of this appendix is to point out the fact that *there exist on D (say on $x^2 + y^2 < 1$) uniformly continuous functions $f(x, y)$ corresponding to which the second derivatives*

$$(7) \qquad \phi_{xx}, \ \phi_{xy} = \phi_{yx}, \ \phi_{yy}$$

of the logarithmic potential (2) *exist on D but* (2) *fails to be of class C^2 in any neighborhood of certain points of D (so that the derivatives* (7) *exist but are not continuous*). In addition, the steps leading to this fact will make it clear that it can be transferred from (1) to (5), if use is made of the considerations applied in [17].

The above-quoted passage in Lebesgue's book is that which, from the unboundedness of the Lebesgue constants of the partial sums of Fourier series, constructs the existence of a continuous function having a divergent Fourier series. But a later passage (Lebesgue, *op. cit.*, pp. 88-89) constructs, on the same basis, a continuous (periodic) function having a Fourier series which is convergent throughout but does not converge uniformly. The statement italicized above follows by using this *second* construction of Lebesgue (instead of appealing, as in [17], to his *first* construction, that in *op. cit.*, pp. 85-87).

It may be mentioned that, while (3) is necessary and sufficient for the existence of $\phi_{xx}(a,b)$, *the uniform existence of the limit* (3) *on all compact subsets of points* (a,b) *of* D *is necessary and sufficient for the existence of a continuous* ϕ_{xx}. In fact, this statement follows from the considerations of [15], pp. 131-133, leading to (3), and from the fact that a function of a single variable has a continuous derivative if and only if the difference quotients have a uniform limit.

Similar remarks hold for the existence and continuity of $\phi_{xy}=\phi_{yx}$ and ϕ_{yy}. The nature of these criteria for the existence and continuity of (7) is one of the reasons why, in contrast to a simple "explicit" example (such as (4)) for the non-existence of (7), a "construction" is needed to exhibit a case in which (7) exist but are not continuous.

These criteria also show that the "construction" can be carried out so as to lead to bounded, but not continuous, functions (7).

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] L. Bianchi, *Opere*, vol. 1_2 (Rome, 1953), pp. 123-135 [1889].

[2] S. S. Chern, "An elementary proof of the existence of isothermal parameters on a surface," (to appear in the Proceedings of the American Mathematical Society).

[3] ———, P. Hartman and A. Wintner, "On isothermic coordinates," *Commentarii Mathematici Helvetici*, vol. 28 (1954), pp. 301-309.

[4] U. Dini, *Serie di Fourier ed altre rappresentazioni analitiche delle funzioni di una variabile reale*, Pisa, 1880.

[5] ———, "Sur la méthode des approximations successives pour les équations aux dérivées partielles du deuxième ordre," *Acta Mathematica*, vol. 25 (1902), pp. 185-230.

[6] P. Hartman and A. Wintner, "On the fundamental equations of differential geometry," *American Journal of Mathematics*, vol. 72 (1950), pp. 757-774.

[7] ——— and A. Wintner, "On the existence of Riemannian manifolds which cannot carry non-constant analytic or harmonic functions in the small," *ibid.*, vol. 75 (1953), pp. 260-276.

[8] O. Hölder, *Beiträge zur Potentialtheorie*, Inaugural-Dissertation (Tübingen), Stuttgart, 1882.

[9] E. Hopf, "Über den funktionalen, inbesondere den analytischen Charakter der Lösungen elliptischer Differentialgleichungen zweiter Ordnung," *Mathematische Zeitschrift*, vol. 34 (1932), pp. 194-233.

[10] A. Korn, "Über Minimalflächen, deren Randkurven wenig von ebenen Kurven abweichen," *Abhandlungen der Königlich Preussischen Akademie der Wissenschaft*, Berlin (1909), Anhang II.

[11] ———, "Zwei Anwendungen der Methode der sukzessiven Annäherungen," *Schwarz Festschrift*, Berlin (1914), pp. 215-229.

[12] M. Lavrentieff, "Sur une classe de représentations continues," *Recueil Mathématique*, vol. 42 (1935), pp. 407-424.

[13] L. Lichtenstein, "Zur Theorie der konformen Abbildung. Konforme Abbildung nicht-analytischer, singularitätenfreier Flächenstücke auf ebene Gebiete," *Bulletin International de l'Académie des Sciences de Cracovie*, ser. A, 1916, pp. 192-217.

[14] ———, "Neuere Entwicklung der Theorie partieller Differentialgleichungen zweiter Ordnung vom elliptischen Typus," *Encyklopädie der mathematischen Wissenschaften*, vol. IIC (1924), pp. 1279-1334.

[15] H. Petrini, "Les dérivées premières et secondes du potentiel logarithmique," *Journal de Mathématiques*, ser. 6, vol. 5 (1909), pp. 127-223.

[16] H. A. Schwarz, *Gesammelte mathematische Abhandlungen*, vol. 2 (1890), pp. 102-107 [1869].

[17] A. Wintner, "On the Hölder restrictions in the theory of partial differential equations," *American Journal of Mathematics*, vol. 75 (1950), pp. 731-738.

ON ALGEBRAIC GROUPS OF TRANSFORMATIONS.*

By André Weil.

In my *Variétés abéliennes* (Hermann, Paris, 1948; quoted hereafter as VA), I gave the rudiments of a theory of algebraic group-varieties. As these have become wholly inadequate to the present state of growth of algebraic geometry, a fuller treatment of this topic will be given here.

To define a group in algebraic geometry, one simply takes over the usual definition and adds the condition that all the objects entering into it must have a meaning from the point of view of the algebraic geometer. This means that the elements of the group must be the points of algebraic varieties in finite number, that the mappings $(x, y) \rightarrow xy$ and $x \rightarrow x^{-1}$ which define the group-structure are mappings in the sense of algebraic geometry, i. e. that their graphs consist of algebraic varieties, and finally that these mappings are everywhere defined in the sense of algebraic geometry. The same can be done in an obvious manner for groups of transformations and for homogeneous spaces.

For simplicity, we consider only groups and spaces consisting of a single variety; this corresponds to the assumption of connectedness in the theory of topological groups. In § I, we shall deal with those properties of groups and transformation-spaces which are birationally invariant, giving what will eventually prove to be a birationally invariant characterization of such spaces; this is obtained by writing down the basic axioms for groups and transformation-spaces at generic points only. Our main purpose is, starting from such objects, to derive from them birationally equivalent objects which are groups and transformation-spaces in the full sense described above. The method which will be followed is very simple, and, I believe, the most natural one which could be imagined; it derives from the observation that the varieties from which one starts, even though they may not be true groups or transformation-spaces, nevertheless contain large pieces or "chunks" (more precisely, open subsets) of such spaces; to isolate these is the purpose of § II. In § III, they are pieced together, by the technique of "abstract varieties," so as to achieve the desired result; the way in which this is done at first

---

* Received December 27, 1954.

requires an enlargement of the groundfield. Then a modified procedure is given whereby our spaces can be exhibited as varieties over the original groundfield; according to an idea which was first applied to similar problems by Matsusaka, and again quite recently by Chow, this is done by building up suitable symmetric products by means of the Chow points of 0-cycles. Some auxiliary results, belonging to the foundations of algebraic geometry, which would have interrupted the treatment of the main topic, are dealt with in an Appendix. A further paper, to appear in this same volume, will contain various applications of the general theory.

§ I. Some birationally invariant results.

The very convenient language of the Zariski topology will be used freely in the following manner. By a *closed subset of a variety* V, or by a *closed set on* V, we shall understand any union of subvarieties of V, *other than* V; thus V itself is not a closed set on V; there is some impropriety in this (the proper words for our concept being "non-dense closed set"), but this seemed preferable to endless repetitions. An *open set on* V is defined as the complement of a closed set on V (this should properly be called a "non-empty open set"). If k is a field of definition for V, we say that a subset of V is *k-closed* if it is closed and its components are algebraic over k, and if moreover it is invariant by all automorphisms over k of the algebraic closure $\bar{k}$ of k. A *k-open set on* V is the complement of a k-closed set.

By a variety, we mean an abstract variety unless the contrary is stipulated. The final models for our groups and transformation-spaces will be constructed in §§ III-IV as abstract varieties; until then, little or nothing would be lost (and nothing would be gained) if we confined our attention to projective or to affine varieties.

1. Let V be a variety, defined over a field k. Let f be a mapping of $V \times V$ into V, defined over k. Consider the following condition on f:

(G1) *If x, y are independent generic points of V over k, and $z = f(x, y)$, then* $k(x,y) = k(x,z) = k(y,z)$.

This is equivalent to saying that $k(x) \subset k(z,y)$ and $k(y) \subset k(x,z)$. It implies that any two of the points x, y, z are independent generic points of V over k and determine the third one uniquely.

Let x, y, t be independent generic points of V over k; (G1) implies that

$(f(x,y),t)$ and $(x,f(y,t))$ are two pairs of independent generic points of V over k. Thus, if (G1) is assumed, the following condition is meaningful:

(G2) *If x, y, t are independent generic points of V over k, then:*

$$f(f(x,y),t)=f(x,f(y,t)).$$

This is of course the associativity condition (but postulated only at independent generic points) for f.

If (G1), (G2) are satisfied, we say that f is a *normal* (internal) *law of composition* on V, and that V, with this law, is a *pre-group*; f will then mostly be written as a multiplication, i. e. as xy instead of $f(x,y)$. If V, with the law f, is a pre-group over k, it is so, a fortiori, over every field K containing k. Let a variety V' be birationally equivalent to V over such a field K; let x,y be independent generic points of V over K, put $z=xy$, and call x',y',z' the generic points of V' over K which correspond to x, y, z respectively. Then $K(x')$, $K(y')$, $K(z')$ are respectively the same as $K(x)$, $K(y)$, $K(z)$, and thus, since $K(z')\subset K(x',y')$, we may write $z'=f'(x',y')$, where f' is a mapping of $V'\times V'$ into V', defined over K. One sees at once that f' satisfies (G1) and (G2); we say that it is the law of composition on V', derived from f by transfer; and we say that V', with the law f', is a pre-group *birationally equivalent* to the pre-group V with the law f. This shows that the concept of a pre-group is invariant under arbitrary birational correspondences; a pre-group can thus be studied on any model, e. g. on an affine or a projective model.

PROPOSITION 1. *Let V be a pre-group, defined over k. There is a uniquely determined mapping ϕ of V into V, which is defined over k and is such that, if we put $s^{-1}=\phi(s)$ for every s on V at which ϕ is defined, the following conditions are fulfilled whenever x, y are independent generic points on V over k:*

(i) $k(x^{-1})=k(x)$; (ii) $(x^{-1})^{-1}=x$; (iii) $y=(x^{-1})(xy)$;

(iv) $x=(xy)(y^{-1})$; (v) $(xy)^{-1}=(y^{-1})(x^{-1})$.

If we put $z=xy$, we have $k(y)\subset k(x,z)$, and therefore there is a mapping λ of $V\times V$ into V, defined over k, such that $y=\lambda(x,z)$; similarly there is a mapping μ of $V\times V$ into V, defined over k, such that $x=\mu(z,y)$. Take t generic on V over $k(x,y,z)$; put $y'=yt$, $z'=zt$; by associativity, the latter relation gives $z'=xy'$. Put $u=\mu(y,z)$; by the definition of μ, this is equivalent to $y=uz$. By (G1), this implies that $k(u,z)=k(y,z)$, and so u, z, t are independent generic points on V over k; therefore, by (G2), we have $(uz)t = u(zt)$, which can be written as $y' = uz'$; as this also shows that u, z' are generic and independent over k, the latter relation implies, by (G1), that $k(u) \subset k(y', z')$. Therefore $k(u)$ is contained in $k(y, z)$, i.e. in $k(x, y)$, and also in $k(y', z')$, i.e. in $k(x, y')$. But, by (G1), y and $y' = yt$ are generic and independent over $k(x)$, and so $k(x, y)$ and $k(x, y')$ are independent regular extensions of $k(x)$; their intersection is therefore $k(x)$, and so we have $k(u) \subset k(x)$, so that we may write $u = \phi(x)$, where ϕ is a mapping of V into V, defined over k. As x, u are no other than $\mu(z, y)$ and $\mu(y, z)$, the relation between them is symmetrical, and we have $x = \phi(u)$ and $k(x) \subset k(u)$, and so $k(x) = k(u)$. We have thus verified (i), (ii), (iii). Also, by (G1), any two of the points x, y, z in $z = xy$ determine the third one uniquely provided they are generic and independent over k; from this it follows that $u = \phi(x)$ is uniquely determined by the relation $y = uz$, and so the function ϕ is uniquely determined by (iii). From now on, write x^{-1} instead of $\phi(x)$.

Let now v be generic on V over $k(x, y)$; put $s = (xy)v = x(yv)$. As x, yv are generic and independent over k, $s = x(yv)$ is equivalent to $yv = x^{-1}s$, by (iii); and again by (iii), this is equivalent to $v = y^{-1}(x^{-1}s)$ since y and yv are generic and independent over k by (G1). By (i), the points x^{-1}, y^{-1} and v are generic and independent over k, and therefore the last relation, by (G2) can be written as $v = (y^{-1}x^{-1})s$. As $s = (xy)v$, and xy, v are generic and independent over k, this shows that $y^{-1}x^{-1} = (xy)^{-1}$, which is (v). This, with $z = xy$, can be written as $z^{-1} = y^{-1}x^{-1}$, which, by (iii), is equivalent to $x^{-1} = y(z^{-1})$; applying (v) to the latter relation, we get $x = (z^{-1})^{-1}(y^{-1})$, which, in view of (ii), gives (iv). This completes the proof.

With the same notations as above, we have $\lambda(x, z) = x^{-1}z$ and $\mu(z, y) = zy^{-1}$. One should observe, however, that the function λ may be defined at a point (s, t) of $V \times V$ without the expression $s^{-1}t$ being defined; in fact, λ may be defined at (s, t) without ϕ being defined at s. A similar remark applies to μ.

COROLLARY. *With the notations of* Prop. 1, *assume that the function* $f(x, y) = xy$ *is defined at* (x^{-1}, x). *Then the function* $x \rightarrow (x^{-1})x$ *is a constant* e, *rational over* k. *If, moreover,* f *is defined at* (e, x), *then* $ex = x$; *if it is defined at* (x, e), *then* $xe = x$.

With x, y and $z = xy$ as before, the assumption implies that $z^{-1}z$, i.e. (by Prop. 1(v)) $(y^{-1}x^{-1})z$ is defined. In the relation $(uv)z = u(vz)$, with u, v generic and independent over $k(z)$, specialize (u, v) to (y^{-1}, x^{-1}) over $k(z)$; by our assumption, the left-hand side is defined; also, u and vz get specialized to y^{-1} and $x^{-1}z$, the latter being defined and equal to y by Prop. 1(iii). Since, by our assumption, f is defined at (y^{-1},y), this gives $z^{-1}z = y^{-1}y$; in other words, the function $x^{-1}x$ has the same value at z and at y. As y, z are independent generic points of V over k, this implies that the function is a constant; as it is defined over k, its constant value must then be rational over k. Putting $e = x^{-1}x$ and replacing x by x^{-1}, we get, in view of Prop. 1(i)-(ii), $e = x(x^{-1})$. Taking t generic over $k(x,y)$, specialize t to x^{-1} in the relation $t(xy) = (tx)y$; the left-hand side becomes y by Prop. 1(iii), and the right-hand side becomes ey provided this is defined. Specializing y to x^{-1} in the same relation, we get $te = t$ provided te is defined.

2. Let V and W be two varieties, defined over a field k. Let f, g be two mappings, both defined over k, of $V \times V$ into V and of $V \times W$ into W, respectively. Consider the following conditions:

(TG1) *For a generic x over k on V, the mapping $u \rightarrow g(x,u)$ of W into W is a birational correspondence between W and W.*

This is equivalent to saying that, if x, u are independent generic points of V and of W, respectively, over k, then $k(x, g(x,u)) = k(x,u)$.

(TG2) *If x, y, u are independent generic points of V, V and W, respectively, over k, then $g(f(x,y),u) = g(x, g(y,u))$.*

If (TG1) is fulfilled, (TG2) is meaningful, since in that case $g(y,u)$ is generic on W over $k(x)$, while $f(x,y)$ is generic on V over $k(u)$ by (G1).

When (G1, 2) and (TG1, 2) are satisfied, we shall say that g is a *normal* (external) *law of composition* on W with respect to the pre-group V, and that W, with this law, is a *pre-transformation space* with respect to V; g will then mostly be written as a multiplication, i. e. as $g(x,u) = xu$; then (TG1), (TG2) appear as $k(x,xu) = k(x,u)$ and $(xy)u = x(yu)$. Just as before, we note that the concept of a normal law is independent of the field of definition, which may be enlarged at will, and that it is birationally invariant; if V' is birationally equivalent to V, and W' to W, the laws f, g can be transferred in an obvious manner to V', W'; the pair V', W', with the laws f', g' obtained from f, g by transfer, is said to be *birationally equivalent* to the pair V, W with the laws f, g. In particular, W, just as V, may be replaced by an affine model.

Take x, y, u as in (TG2); put $z = xy$, $v = yu$. By (TG1), v is generic on W over $k(x)$; so is xv, again by (TG1); therefore $x^{-1}(xv)$ is defined. But (TG2) can be written as $zu = xv$, and so we have $x^{-1}(zu) = x^{-1}(xv)$.

As x^{-1}, z and u are independent generic points of V, V and W over k, we can apply (TG2) to the left-hand side, which is therefore equal to $(x^{-1}z)u$, i. e. to yu by Prop. 1 (iii), i. e. to v. This proves $x^{-1}(xv) = v$; as x, v are independent generic points of V, W over k, this must therefore remain true for any pair of such points.

The conditions stated above may be strengthened by assuming "generic transitivity," which means the following condition:

(H). *If x, u are independent generic points of V and of W, respectively, over k, then $g(x,u)$ is a generic point of W over $k(u)$.*

In that case, we say that W is a *pre-homogeneous space* with respect to the pre-group V. This condition is equivalent to saying that the graph of g on $V \times W \times W$ has the projection $W \times W$ on $W \times W$ (in the sense of my *Foundations*; the set-theoretic projection then contains a open subset of $W \times W$, by Prop. 10 of the Appendix).

3. The following result shows that a normal law of composition may be obtained from a mapping satisfying much weaker conditions than those stated above.

PROPOSITION 2. *Let V, W be two varieties, defined over a field k. Let g be a mapping of $V \times W$ into W, defined over k, satisfying* (TG1) *and the following condition:*

(TG2′) *There are two independent generic points x, y of V over k and a generic point z of V over k such that $g(z,u) = g(x, g(y,u))$ for u generic on W over $k(x,y,z)$.*

Let $k(\bar{x})$ be the smallest field of definition containing k for the mapping $u \rightarrow g(x,u)$ of W into W, $\bar{x}$ being a generic point over k of a variety $\bar{V}$. Then one can write $\bar{x} = \phi(x)$ and $g(x,u) = \bar{g}(\bar{x},u)$, where ϕ is a mapping from V to $\bar{V}$ and $\bar{g}$ a mapping from $\bar{V} \times W$ to W, both defined over k; putting $\bar{y} = \phi(y)$, $\bar{z} = \phi(z)$, we have $k(\bar{z}) \subset k(\bar{x},\bar{y})$ and may write $\bar{z} = \bar{f}(\bar{x},\bar{y})$, where $\bar{f}$ is a mapping from $\bar{V} \times \bar{V}$ to $\bar{V}$, defined over k. Finally, $\bar{f}$ and $\bar{g}$ satisfy the conditions (G1, 2), (TG1, 2) *and define $\bar{V}$ as a pre-group and W as a pre-transformation space with respect to $\bar{V}$.*

In the first place, the smallest field of definition containing k for $u \rightarrow g(x,u)$ is contained in $k(x)$ and is therefore, by Prop. 3 of the Appendix, a finitely generated regular extension of k; this may always be written as $k(\bar{x})$, e. g. by taking $\bar{x}$ as a suitable point in an affine space, which has then a locus $\bar{V}$ over k and may be written as $\phi(x)$. As $g(x,u)$ is then rational over $k(\bar{x},u)$, it may be written as $\bar{g}(\bar{x},u)$, or more briefly as $\bar{x}u$; (TG2′) can then be written as $g(z,u)=\bar{x}(\bar{y}u)$, which shows that the function $u \rightarrow g(z,u)$ is defined over $k(\bar{x},\bar{y})$, so that $k(\bar{z}) \subset k(\bar{x},\bar{y})$; we may then write $\bar{z}=\bar{f}(\bar{x},\bar{y})$, or more briefly $\bar{z}=\bar{x}\bar{y}$. It is clear that $\bar{g}$, $\bar{f}$ satisfy (TG1, 2); we have to show that $\bar{f}$ satisfies (G1, 2). By (TG1), if $v=g(x,u)$, the mapping $u \rightarrow v$ is a birational correspondence between W and W, defined over $k(\bar{x})$; its inverse must then be defined over the same field, so that we have $k(u) \subset k(\bar{x},v)$ and may write $u=h(\bar{x},v)$. Notations being as before, put $w=g(y,u)$; as this, by (TG1), is generic over $k(x)$ on W, the relation in (TG2′), which can be written as $\bar{z}u=\bar{x}w$, is equivalent to $w=h(\bar{x},\bar{z}u)$. This shows that the mapping $u \rightarrow w$ is defined over $k(\bar{x},\bar{z})$; since its smallest field of definition is $k(\bar{y})$, we get $k(\bar{y}) \subset k(\bar{x},\bar{z})$. Similarly, we have $w=\bar{y}u$ and therefore $u=h(\bar{y},w)$; then the relation in (TG2′) can be written as $\bar{x}w=\bar{z}h(\bar{y},w)$, from which we conclude in the same manner that $k(\bar{x}) \subset k(\bar{z},\bar{y})$. This shows that f satisfies (G1).

Now, if $\bar{x}_1$, $\bar{x}_2$ are any two generic points of $\bar{V}$ over k, there is an isomorphism σ of $k(\bar{x}_1)$ onto $k(\bar{x}_2)$ over k which maps $\bar{x}_1$ onto $\bar{x}_2$. Take u generic on W over $k(\bar{x}_1,\bar{x}_2)$, and put $u_1=\bar{x}_1u$, $u_2=\bar{x}_2u$. Then σ maps the graph of $u \rightarrow u_1$ onto the graph of $u \rightarrow u_2$. If $u_1=u_2$, these two functions coincide, and therefore, by Prop. 4 of F-IV$_2$, σ must induce the identity on the smallest field of definition of the first function. As this field is $k(\bar{x}_1)$, we have thus shown that $u_1=u_2$ implies $\bar{x}_1=\bar{x}_2$. Now let $\bar{x}$, $\bar{y}$, $\bar{t}$, u be independent generic points over k on $\bar{V}$, $\bar{V}$, $\bar{V}$, W; put $\bar{x}_1=(\bar{x}\bar{y})\bar{t}$ and $\bar{x}_2=\bar{x}(\bar{y}\bar{t})$, these being defined because $\bar{f}$ satisfies (G1). We have to show that $\bar{x}_1=\bar{x}_2$; by (G1), they are both generic over k on $\bar{V}$, and u is generic over $k(\bar{x}_1,\bar{x}_2)$ on W, so that we need only show that $\bar{x}_1u=\bar{x}_2u$. By (TG1), $\bar{x}(\bar{y}(\bar{t}u))$ is defined; by (TG2), this is the same as $(\bar{x}\bar{y})(\bar{t}u)$, which, again by (TG2), is the same as $\bar{x}_1u$ since $\bar{x}\bar{y}$, $\bar{t}$, u are independent generic points of $\bar{V}$, $\bar{V}$, W over k by (G1). Similarly $\bar{x}(\bar{y}(\bar{t}u))$ is the same as $\bar{x}((\bar{y}\bar{t})u)$ by (TG2), and this is the same as $\bar{x}_2u$ by (TG2) and (G1). This concludes the proof.

The external normal law of composition $\bar{g}$ constructed in Prop. 2 satisfies, in addition to (TG1, 2), the following condition:

(TG3) *If x is generic over k on V, $k(x)$ is the smallest field of definition containing k for the mapping $u \rightarrow g(x,u)$ of W into W.*

Whenever (TG3) is satisfied in addition to (G1, 2) (TG1, 2), we will say that V operates *faithfully* on W by g.

If V, W and a mapping g of $V \times W$ into W are given, and g satisfies (TG1, 2′, 3), then Prop. 2 shows that $(x, y) \rightarrow z$ (where x, y, z are the points of V which appear in (TG2′)) is a normal internal law of composition on V and that g is a normal external law with respect to the pre-group defined by f on V. In particular, if V, W, f and g are given and f, g satisfy (TG1, 2, 3), then f satisfies (G1, 2).

§ II. Construction of chunks.

4. Let V, W be a pre-group and a pre-transformation space and f, g the internal and external normal laws belonging to them, these being all defined over a field k; we will mostly write f, g multiplicatively, as has already been done in § I. Instead of saying that f is defined at a point (s, t) of $V \times V$, we shall frequently say that st is defined; similarly, when we say for instance that $s^{-1}((st)a)$ is defined, for s, t on V and a on W, this will mean the following: (i) f is defined at (s, t), with the value st; (ii) g is defined at (st, a), with the value $(st)a$; (iii) $x \rightarrow x^{-1}$ is defined at s, with the value s^{-1}; (iv) g is defined at $(s^{-1}, (st)a)$, with the value $s^{-1}((st)a)$. We recall that two expressions, built up from functions which are defined over k, coincide for all values of the variables for which they are both defined provided they are defined and coincide when the variables are given independent generic values over k. This applies for instance to the formulas in (G2) and (TG2) which express the associativity of f and g.

We say that V is a *group-variety* or a *group* if f is everywhere defined on $V \times V$ and $x \rightarrow x^{-1}$ is everywhere defined on V; then the corollary of Prop. 1 shows that there is a neutral element e on V with the usual properties. If V is a group, W will be called a *transformation-space* with respect to V if g is everywhere defined on $V \times W$; if, moreover, V operates transitively on W in the usual sense, i. e. if to every pair a, b on W there is an $s \in V$ such that $b = sa$, then W is called a *homogeneous space* with respect to V.

In § III, it will be shown that, to every pre-group V, there is a birationally equivalent group V', and that, to every pre-transformation space W with respect to V, there is a birationally equivalent transformation-space W' with respect to V'. The proof of this will include a proof of the fact that W is biregularly equivalent to an open subset of W' if and only if it fulfills the following condition:

(C) *If a is any point of W, and x a generic point of V over $k(a)$, then xa and $x^{-1}(xa)$ are defined.*

A pre-transformation space W with respect to V which fulfills this condition will be called a *chunk of transformation-space,* or more briefly a *chunk.*

For similar reasons, if W is a pre-homogeneous space, i. e. if g satisfies (H), we say that W is a *homogeneous chunk* if it satisfies (C) and the following:

(HC) *If a and x are as in* (C), *xa is generic over $k(a)$ on W.*

Finally, V itself will be called a *group-chunk* if it is a homogeneous chunk with respect to left-translations and if x^{-1} is everywhere defined on it, or in other words if it satisfies the following:

(GC1) *If s is any point on V, and x a generic point of V over $k(s)$, then xs and $x^{-1}(xs)$ are defined and xs is generic over $k(s)$ on V.*

(GC2) *For every s on V, s^{-1} is defined.*

PROPOSITION 3. *Call Ω the set of those points a on W such that xa and $x^{-1}(xa)$ are defined for x generic over $k(a)$ on V. Then Ω is a k-open subset of W; Ω and all k-open subsets of Ω are chunks; if $a \in \Omega$, we have $x^{-1}(xa) = a$, $k(x,a) = k(x,xa)$, and a is a point of the locus of xa over $k(a)$ on W.*

Call F the set of points on $V \times W$ where g is not defined; by Prop. 8 of the Appendix, this is a k-closed subset of $V \times W$. Let Γ be the graph of the mapping $(x,u) \rightarrow x^{-1}u$ of $V \times W$ into W, i. e. the locus of $(x,u,x^{-1}u)$ over k for x, u generic and independent over k on V, W. Call F' the k-closed subset of $V \times W \times W$ consisting of all points (x,u,v) with $(x,v) \in F$; let F'' be the union of the projections of the components of $\Gamma \cap F'$ on the product of the first two factors of $V \times W \times W$ (this being understood as in F-IV$_3$ and F-VII$_3$; F'' is the closure, in the Zariski topology, of the set-theoretic projection of $\Gamma \cap F'$ on $V \times W$; cf. Appendix, Prop. 10). It will now be shown that Ω is the same as the set Ω_1 of the points a on W such that $V \times a$ is not contained in $F \cup F''$. In fact, for xa to be defined, it is necessary and sufficient that $V \times a$ should not be contained in F; let Ω_0 be the set of points a with this property; it contains both Ω and Ω_1. If $a \in \Omega_0 - \Omega$, $x^{-1}(xa)$ is not defined; as x^{-1} is generic over $k(a)$ at the same time as x, this is equivalent to saying that $x(x^{-1}a)$ is not defined; as $x^{-1}a$ is defined, the point $(x,a,x^{-1}a)$ is then in $\Gamma \cap F'$, and therefore (x,a) is in F'', so that $V \times a \subset F''$ and $a \notin \Omega_1$. Conversely, if $a \in \Omega_0 - \Omega_1$, then (x,a) is in the projection of one of the components of $\Gamma \cap F'$, and so, if (y,u,v) is a generic point of that component over $\bar{k}$, (x,a) is a specialization of (y,u) over $\bar{k}$. As (y,u,v) is on Γ and $x^{-1}a$ is defined, v has then the unique specialization $x^{-1}a$ over $(y,u)\rightarrow(x,a)$ with respect to $\bar{k}$; therefore $(x,a,x^{-1}a)$ is in F', and so $x(x^{-1}a)$ is not defined and a is not in Ω. This proves that $\Omega=\Omega_1$; the latter set being k-open by Prop. 7 of the Appendix, Ω is k-open.

As $x^{-1}(xu)=u$ for x, u generic and independent over k on V, W, we must have $x^{-1}(xa)=a$ whenever the left-hand side is defined, and so for $a\epsilon\Omega$ and x generic over $k(a)$; this implies that $k(a)\subset k(x,xa)$, so that $k(x,a)=k(x,xa)$. Let y be generic over $k(x,a)$ on V; then (x^{-1},xa) is a specialization of (y,xa) over $k(x,a)$; as the former point is not in F, and F is k-closed, (y,xa) is not in F, and so $y(xa)$ is defined. As yx is generic over $k(a)$ by (G1) and is therefore a generic specialization of x over $k(a)$, $(yx)a$ is defined and is a generic specialization of xa over $k(a)$. By associativity, we have $y(xa)=(yx)a$ since both sides are defined; as $x^{-1}(xa)$ is defined, it is a specialization of $y(xa)$, and therefore also of $(yx)a$ and of xa, over $k(a)$. This shows that a is a specialization of xa over $k(a)$, i.e. that it is a point of the locus of xa over $k(a)$ on W. If now Ω' is any k-open subset of Ω, then the set $C=W-\Omega'$ is k-closed, and so, if xa is in C, a must be in C; in other words, if a is in Ω', so is xa; it is then clear that Ω' is a chunk.

The locus of xa over $k(a)$ could be described as the closure of the orbit of a under V on W.

COROLLARY. *Notations being as in Prop.* 3, *call* Ω_h *the set of the points* a *of* Ω *such that* W *is the locus of* xa *over* $k(a)$. *Then* Ω_h *is* k*-open or empty according as* W *is pre-homogeneous or not. In the former case,* Ω_h *and all* k*-open subsets of* Ω_h *are homogeneous chunks; and, if* a, b *are any two points of* Ω_h, *there are two generic points* x, y *of* V *over* $k(a,b)$ *such that* $xa=yb$.

Except for the last assertion, this is an immediate consequence of Prop. 11 of the Appendix, applied to the k-open set Ω of Prop. 3. Let now a, b be in Ω_h; take x, y generic on V over $k(a,b)$, and put $u=xa$, $v=yb$. Then the loci of u and of v over $k(a,b)$ are W, and so there is an isomorphism of $k(a,b,u)$ onto $k(a,b,v)$ over $k(a,b)$ which maps u onto v; this can be extended to an isomorphism σ of $k(a,b,x)$ onto some extension of $k(a,b,v)$; then x^{σ} is generic on V over $k(a,b)$ and we have $u^{\sigma}=x^{\sigma}a$, i.e. $x^{\sigma}a=yb$, so that x^{σ} and y satisfy the conditions stated in the corollary.

Finally, in order to construct a group-chunk from a given pre-group V, one need only observe that the graph V_1 of the function $x\rightarrow x^{-1}$ is a subvariety of $V\times V$, birationally equivalent to V, and that, if we transfer that function to V_1, we get an everywhere biregular birational correspondence between V_1 and itself since it is the same as the function induced on V_1 by the mapping $(x,y) \rightarrow (y,x)$ of $V \times V$ onto itself. Therefore we may assume that we have started from a pre-group V on which x^{-1} was everywhere defined; had that not been the case, one would merely have had to replace V by V_1 to make it so. Call now Ω_h the set of the points $s \in V$ with the property stated in (GC1); by the corollary of Prop. 3, this is a k-open subset of V; as $x \rightarrow x^{-1}$ is an everywhere biregular mapping of V onto itself, it transforms Ω_h into a k-open set Ω_h^{-1}; then $\Omega_h \cap \Omega_h^{-1}$ is a group-chunk.

Thus we have constructed chunks for the three kinds of objects under consideration, viz., transformation-spaces, homogeneous spaces and groups. If W is a pre-transformation space, defined over k, the set W' of simple points on W is a k-open set on W; by applying Prop. 3 to W', we obtain a non-singular chunk. Similarly, one would get an everywhere normal chunk by taking for W' the set of points where W is normal, this being k-open by Corollary 3 of Prop. 8 of the Appendix. It will presently be seen that homogeneous chunks and in particular group-chunks are always non-singular, so that no special procedure is required to make them such.

By Corollary 2 of Prop. 8 of the Appendix, if one has constructed a chunk, one can at once derive from it a birationally equivalent chunk which is an affine variety; this also applies to homogeneous chunks. As to group-chunks, starting from a pre-group which we take to be an affine variety, and replacing it by the graph of the function x^{-1} on it, we get for our pre-group an affine model V on which x^{-1} is everywhere defined. Let V' be a k-open set on V which is a homogeneous chunk; let $x = (x_1, \cdots, x_m)$ be a generic point of V over k; take a polynomial P with coefficients in k which is 0 on $V - V'$ but not on V; as x^{-1} and $P(x)$ are everywhere defined functions on V, so is $P(x^{-1})$. Call V'' the locus of

$$(x_1, \cdots, x_m, 1/P(x), 1/P(x^{-1}))$$

over k in affine space; this is biregularly equivalent to the k-open subset determined on V by the inequalities $P(x) \neq 0$, $P(x^{-1}) \neq 0$. This is a group-chunk. We have thus proved the following:

PROPOSITION 4. *To every pre-homogeneous space* (resp. *pre-group*) *defined over k, there is a birationally equivalent homogeneous chunk* (resp. *group-chunk*) *which is an affine variety, defined over k. To every pre-transformation space W and every point a on W with the property stated in* (C), *there is a birationally equivalent chunk W' which is an affine variety and is such that the birational correspondence between W and W' is biregular at a; if a is simple on W, W' may be taken non-singular; if W is normal at a, W' may be taken everywhere normal.*

5. Proposition 5. *Let V be a group-chunk and W a pre-transformation space with respect to V; let k be a field of definition for V, W. Then, if s is any point of V, and u a generic point of W over $k(s)$, su and $s^{-1}(su)$ are defined; the mapping $u \rightarrow su$ is a birational correspondence between W and itself; and $k(s,u) = k(s,su)$.*

Take x, y generic and independent on V over $k(s,u)$; put $y' = yx^{-1}$ and $u' = xu$; by (G1) and (TG1), y' and u' are generic and independent over k on V, W, so that $y'u'$ is defined; as we have shown $x^{-1}u'$ to be defined and equal to u, we get, by associativity, $y'u' = yu$. We now show that the expression obtained by substituting s for y in $y'u'$, i. e. in $(yx^{-1})(xu)$, is defined. In fact, since V is a group-chunk, the mapping $(z,t) \rightarrow t^{-1}z^{-1}$, where z, t are generic and independent over $k(s)$ on V, is defined at (s,t), and its value $t^{-1}s^{-1}$ at that point is generic over $k(s)$ on V; this implies that the mapping $(z,t) \rightarrow (t^{-1}z^{-1})^{-1}$ is also defined there; as this is only another expression for the mapping $(z,t) \rightarrow zt$, we conclude that the latter is defined at (s,t), i. e. that st is defined, and that st is generic over $k(s)$; substituting x^{-1} for t, this shows that $x' = sx^{-1}$ is defined and generic over $k(s)$ on V, and a fortiori that the mapping $y \rightarrow yx^{-1}$ of V into V is defined at s, with the value x'. The mapping $u \rightarrow u'$ is defined at u, with the value u' which is generic over $k(x,s)$ on W by (TG1). So x' and u' are generic and independent over k on V, W, and $x'u'$ is defined; more precisely, we have shown that the mapping $(y,u) \rightarrow y'u' = (yx^{-1})(xu)$ of $V \times W$ into W, which is defined over $k(x)$, is defined at (s,u). As this is only another expression for the mapping $(y,u) \rightarrow yu$, this implies that the latter is defined at (s,u), i. e. that su is defined, and that these mappings have the same value there, i. e. that $x'u' = su$. By (TG1), $x'u'$ is generic on W over $k(x,s)$; therefore su is generic on W over $k(s)$. But then our assumptions on s, u are also satisfied by s^{-1}, su, so that it follows from what we have already proved that $s^{-1}(su)$ is defined; its value must then be u, since $x^{-1}(xu) = u$, and so we have $k(u) \subset k(s,su)$, and therefore $k(s,u) = k(s,su)$; this means that $u \rightarrow su$ is a birational correspondence between W and W.

Corollary. *Assumptions being as in Prop. 5, assume also that V operates faithfully on W; let s, s' be any two points of V, and let u be generic on W over $k(s,s')$. Then $su = s'u$ implies $s = s'$.*

Take x generic over $k(s,s',u)$ on V. Since V is a group-chunk, xs is defined and generic over $k(u)$ on V; and su is defined and generic over $k(x)$ on W by Prop. 5; by associativity, this gives $(xs)u = x(su)$. Similarly we have $(xs')u = x(s'u)$. Therefore $su = s'u$ implies $(xs)u = (xs')u$. But then we can repeat the argument used in the proof of Prop. 2; there is an isomorphism σ of $k(xs)$ onto $k(xs')$, mapping xs onto xs'; as this transforms the graph of the function $u \rightarrow (xs)u$ into itself, and as $k(xs)$ is the smallest field of definition for this graph because of the assumption of faithfulness, σ must be the identity, and $xs = xs'$. As $s = x^{-1}(xs)$ and $s' = x^{-1}(xs')$, this gives $s = s'$.

PROPOSITION 6. *Let V be a group-chunk and W a chunk of transformation-space with respect to V. Let s be any point on V and (a, b) any point on the graph of the birational correspondence $u \rightarrow su$ between W and itself; then the latter is biregular at (a, b), sa and $s^{-1}b$ are defined, and we have $sa = b$, $s^{-1}b = a$.*

We first show that sa is defined. Take x, y, u generic and independent on V, V, W over $k(a, b, s)$; and consider the mapping $(x, u) \rightarrow y^{-1}((yx)u)$, defined over $k(y)$, of $V \times W$ into W. By (GC1), $x \rightarrow yx$ is defined at s, with a value ys which is generic on V over $k(s, a)$. By (C), the mapping $(x, u) \rightarrow xu$ is defined at (ys, a), and so the mapping $(x, u) \rightarrow (yx)u$ is defined at (s, a), with the value $(ys)a$. At the same time, $(ys)u$ is defined since u is generic on W over $k(y, s)$, and $y(su)$ is defined because su is defined and generic over $k(y)$ by Prop. 5; by associativity, this gives $(ys)u = y(su)$. As (a, b) is on the graph of $u \rightarrow su$, and (u, su) is a generic point of that graph over $k(y, s)$, (a, b) is a specialization of (u, su) over $k(y, s)$; but then the relation $(ys)u = y(su)$ implies $(ys)a = yb$. By Prop. 3, the mapping $v \rightarrow y^{-1}v$ is defined at yb, i. e. at $(ys)a$, with the value b. We have thus proved that $(x, u) \rightarrow y^{-1}((yx)u)$ is defined at (s, a), with the value b; as this is but an expression for $(x, u) \rightarrow xu$, this shows that sa is defined and equal to b. Interchanging a, s with b, s^{-1}, and making use of Prop. 5, we see from this that $s^{-1}b$ is defined and equal to a. This implies a fortiori that the mappings $u \rightarrow su$, $u \rightarrow s^{-1}u$ are respectively defined at a, b, with values b, a; this means that the birational correspondence $u \rightarrow su$ is biregular at (a, b).

COROLLARY. *Every homogeneous chunk is non-singular.*

Let W be such a chunk with respect to a pre-group V; replace V by a birationally equivalent group-chunk. For any a on W, take x generic on V over $k(a)$; then $u \rightarrow xu$ is a birational correspondence between W and itself, transforming a into the generic point xa of W over $k(a)$, and biregular at a. As xa is simple on W, a must therefore be simple on W.

§ III. Construction of spaces.

From now on, until the end of § III, *V and W will denote respectively a group-chunk and a chunk of transformation-space with respect to V, both being at the same time assumed to be affine varieties; k will denote a common field of definition for V and W and for the normal laws given on them.*

6. Let n, n' be the dimensions of V, W, and take $N > 4n$ and also $> 3n + n'$; take N independent generic points $t_1, \cdots, t_N$ over k on V; put $K = k(t_1, \cdots, t_N)$. Let u be a generic point of W over K; put $S_\alpha = W$ and $u_\alpha = t_\alpha u$ for $1 \leqq \alpha \leqq N$. Take the u_α as the corresponding generic points of the varieties S_α over K; this defines birational correspondences $T_{\beta\alpha}$ between any two of the S_α; as we may write, by associativity, $u_\beta = (t_\beta t_\alpha^{-1}) u_\alpha$, $T_{\beta\alpha}$ is the birational correspondence $u \rightarrow (t_\beta t_\alpha^{-1}) u$ between W and itself. Proposition 6 of § II shows that $T_{\beta\alpha}$ is biregular at any pair of points on its graph. Therefore the varieties S_α (with empty "frontiers") and the $T_{\beta\alpha}$ may be used to define an abstract variety $\boldsymbol{S}$. Call $\bar{u}$ the generic point of $\boldsymbol{S}$ over K with the representatives u_α and write $\bar{u} = \Phi(u)$, $u = \Psi(\bar{u})$; Φ is a birational correspondence between W and $\boldsymbol{S}$, and Ψ is its inverse; both are defined over K. Let a be any point of W; by Prop. 4 of the Appendix, there is an α such that t_α is generic over $k(a)$; as W is a chunk, $t_\alpha a$ is then defined; this means that Φ is defined at a, $\Phi(a)$ being the point of $\boldsymbol{S}$ with the representative $t_\alpha a$ on S_α. As $t_\alpha^{-1}(t_\alpha a)$ is then defined and has the value a, Ψ is defined at the point $\Phi(a)$ with the value a. This shows that Φ is a biregular mapping of W onto its set-theoretic image $\Phi(W)$ on $\boldsymbol{S}$; as the latter is the set of points of $\boldsymbol{S}$ where Ψ is defined, it is K-open on $\boldsymbol{S}$ by Prop. 8 of the Appendix. Once and for all, we will agree to denote by $\bar{a}$ the image $\Phi(a)$ of $a \in W$ by Φ in $\Phi(W)$.

All this can be applied to the case when W is taken to be the same as V, V acting upon itself by left-translations. Let $\boldsymbol{G}$ be the abstract variety thus obtained from V; call Φ_0 the birational correspondence between V and $\boldsymbol{G}$ which takes the place of the mapping Φ defined above; and call Ψ_0 its inverse. We transfer to $\boldsymbol{G}$ the normal law on V by means of Φ_0; in other words, for x, y generic and independent on V over K and $\bar{x} = \Phi_0(x)$, $\bar{y} = \Phi_0(y)$, we define $\bar{x}\bar{y} = \Phi_0(xy)$, which implies $\bar{x}^{-1} = \Phi_0(x^{-1})$, and prove that this makes $\boldsymbol{G}$ into a group. In fact, the representative of $\bar{x}^{-1}$ on G_β is $t_\beta x^{-1} = (t_\beta x_\alpha^{-1}) t_\alpha$; if $\bar{s}$ is a point of $\boldsymbol{G}$ with a representative s_α on G_α, we can choose β such that t_β is generic on V over $k(s_\alpha, t_\alpha)$. Then, since V is a group-chunk, $t_\beta s_\alpha^{-1}$ is defined and generic on V over $k(s_\alpha, t_\alpha)$, and so $x \rightarrow (t_\beta x^{-1}) t_\alpha$ is defined at s_α; this means that $\bar{s}^{-1}$ is defined and has a representative on G_β. Similarly, if we write $t = t_\gamma t_\alpha^{-1}$, the representative of $\bar{x}\bar{y}$ on G_γ is $t_\gamma xy = ((tx_\alpha)t_\beta^{-1})y_\beta$; let $\bar{r}$, $\bar{s}$ be two points of $\boldsymbol{G}$ with representatives r_α, s_β on G_α, G_β respectively; by Prop. 4 of the Appendix, we can choose γ so that t_γ is generic on V over $k(r_\alpha, s_\beta, t_\alpha, t_\beta)$; the same will then be true of t, and also of tr_α and of $(tr_\alpha)t_\beta^{-1}$ since V is a group-chunk; for a similar reason, this implies that $(x, y) \rightarrow ((tx)t_\beta^{-1})y$ is defined at (r_α, s_β), and this completes the proof that $\boldsymbol{G}$ is a group.

Now, going back to the space $\boldsymbol{S}$ constructed before, we transfer to $\boldsymbol{G}$, $\boldsymbol{S}$, by means of the birational correspondences Φ_0, Φ, the normal law given for V, W; in other words, for x, u generic and independent over K on V, W, and for $\bar{x} = \Phi_0(x)$, $\bar{u} = \Phi(u)$, we define $\bar{x}\bar{u} = \Phi(xu)$, and prove that this makes $\boldsymbol{S}$ into a transformation-space with respect to $\boldsymbol{G}$. In fact, the representative of $\bar{x}\bar{u}$ on S_γ is $((tx_\alpha)t_\beta^{-1})u_\beta$, where $t = t_\gamma t_\alpha^{-1}$ as before; the rest of the proof is then quite similar to the proof given above.

Naturally, if W is non-singular, $\boldsymbol{S}$ is non-singular; if W is everywhere normal, $\boldsymbol{S}$ is everywhere normal. Finally, if W is a homogeneous chunk, $\boldsymbol{S}$ is a homogeneous space. In fact, in that case, let $\bar{a}$, $\bar{b}$ be any two points of $\boldsymbol{S}$, with representatives a_α, b_β in S_α, S_β respectively. Take x generic over $K(\bar{a}, \bar{b})$ on V; put $\bar{x}' = \bar{x}\bar{a}$, $\bar{x}'' = \bar{x}\bar{b}$. For u generic over $K(x)$ on W, we have $\Psi(\bar{x}\bar{u}) = (xt_\alpha^{-1})u_\alpha$; as W is a homogeneous chunk, $x' = (xt_\alpha^{-1})a_\alpha$ is defined and generic over $K(\bar{a}, \bar{b})$ on W, and therefore we have $x' = \Psi(\bar{x}')$; similarly we have $x'' = \Psi(\bar{x}'')$ with $x'' = (xt_\beta^{-1})b_\beta$ generic over $K(\bar{a}, \bar{b})$ on W. That being so, there is an isomorphism of $K(\bar{a}, \bar{b}, x')$ onto $K(\bar{a}, \bar{b}, x'')$ over $K(\bar{a}, \bar{b})$ which maps x' onto x''; this can be extended to an isomorphism σ of $K(\bar{a}, \bar{b}, \bar{x})$ onto some extension of $K(\bar{a}, \bar{b}, x'')$. Then we have $\bar{x}^\sigma\bar{a} = \bar{x}'' = \bar{x}\bar{b}$, and so $\bar{b} = \bar{x}^{-1}\bar{x}^\sigma\bar{a}$.

7. From now on, it will be assumed that W and consequently $\boldsymbol{S}$ are everywhere normal. With this assumption, we shall construct an abstract variety $\boldsymbol{S}'$, defined over k, and a birational correspondence F between $\boldsymbol{S}'$ and W, also defined over k, so that the birational correspondence $\Phi \circ F$, defined over K, between $\boldsymbol{S}'$ and $\boldsymbol{S}$ is an everywhere biregular mapping of $\boldsymbol{S}'$ onto $\boldsymbol{S}$. This construction can then be applied to V itself, giving a variety $\boldsymbol{G}'$ and a birational correspondence F_0 between $\boldsymbol{G}'$ and V, both defined over k, such that $\Phi_0 \circ F_0$ is biregular between $\boldsymbol{G}'$ and $\boldsymbol{G}$. Transferring the normal laws for V, W to $\boldsymbol{G}'$, $\boldsymbol{S}'$ by means of F, F_0, we see that we have thus constructed a group $\boldsymbol{G}'$ and a transformation-space $\boldsymbol{S}'$, birationally equivalent to V, W over k; if W is pre-homogeneous and we have constructed $\boldsymbol{S}$ as a homogeneous space, $\boldsymbol{S}'$ will be a homogeneous space.

In constructing $\boldsymbol{S}'$, we may assume that V operates faithfully on W; in fact, if this were not so, one could replace V by another pre-group $\bar{V}$ satisfying this condition, according to Prop. 2 of § I, no. 3.

Notations will now be the same as in no. 6, with the additional assumptions that W and consequently $\boldsymbol{S}$ are everywhere normal, and that V acts faithfully on W, so that $\boldsymbol{G}$ acts faithfully on $\boldsymbol{S}$.

Let k' be any field containing k. Let $\sum_{i=1}^{r}(s_i)$ be a cycle of dimension 0 on V, rational over k', and assume that $s_i \neq s_j$ whenever $i \neq j$. Then, if we put $k'' = k'(s_1, \cdots, s_r)$, k'' is a Galois extension of k', i. e. separably algebraic and normal over k'. Call K'' the compositum of K and k''; let u be a generic point of W over K'', and put $w_i = s_i u$. If m is the dimension of the ambient affine space to W, we write $w_i = (w_{i1}, \cdots, w_{im})$. Put now

$$y(T, U) = \prod_{i=1}^{r} \left(T - \sum_{\mu=1}^{m} w_{i\mu} U_\mu\right) \tag{1}$$

where $T, U_1, \cdots, U_m$ are indeterminates; let y be the point, in an affine space of suitable dimension, whose coordinates are all the coefficients of the homogeneous polynomial $y(T, U)$ except that of T^r; this is the so-called "Chow point" of the cycle $\sum_i (w_i)$, and $y(T, U)$ is its "Chow form."

As V acts faithfully on W, and the s_i have been assumed to be distinct, the corollary of Prop. 5, § II, no. 5, shows that the w_i are all distinct. We can therefore apply to them the following general result:

LEMMA. *If in* (1) *we take the w_i to be any set of distinct points, and k_0 is the prime field, then the w_i are separably algebraic over $k_0(y)$.*

By F-I$_5$, Th. 1, we need only show that a derivation D of the field $k_0(w_1, \cdots, w_r)$ over $k_0(y)$ must be trivial. In fact, applying D to (1), we get:

$$0 = \sum_{i=1}^{r} \left(T - \sum_{\mu} w_{i\mu} U_\mu\right)^{-1} \sum_{\mu} D w_{i\mu} U_\mu;$$

as the w_i are all distinct, this cannot be an identity in $T, U_1, \cdots, U_m$ unless all the $Dw_{i\mu}$ are 0.

PROPOSITION 7. *Notations being as defined above, we have $k'(y) = k'(u)$ provided the s_i are all distinct and satisfy the following condition:*

(S) *The set of points $\bar{s}_i = \Phi_0(s_i)$ on $\boldsymbol{G}$ is not mapped onto itself by any right-translation.*

The cycle $\sum(w_i)$ is the image of the cycle $\sum(s_i)$ by the mapping

$x \rightarrow xu$ of V into W; it is therefore rational over $k'(u)$. By the main theorem on symmetric functions (VA, no. 7, Th. 1), this implies that y is rational over $k'(u)$, i. e. that $k'(y) \subset k'(u)$. On the other hand, the lemma shows that the w_i are separably algebraic over $k''(y)$; as we have $u = s_i^{-1}w_i$ by Prop. 5 of § II, no. 5, u is therefore separably algebraic over $k''(y)$, hence also over $k'(y)$. Let σ be any automorphism over $k'(y)$ of the algebraic closure of $k'(y)$; as it induces an isomorphism of $k'(u)$ onto $k'(u^\sigma)$ over k', u^σ is generic on W over k', so that $s_i u^\sigma$ is defined by Prop. 5. This gives $(s_i u)^\sigma = s_i^\sigma u^\sigma$, i. e. $w_i^\sigma = s_i^\sigma u^\sigma$. But the decomposition of the homogeneous polynomial $y(T, U)$ into linear factors is uniquely determined; applying σ to (1), we see thus that the w_i^σ must be the same as the w_i except for a permutation, i. e. that there is a permutation $i \rightarrow \sigma(i)$ such that $w_i^\sigma = w_{\sigma(i)}$. This can be written as $s_i^\sigma u^\sigma = s_{\sigma(i)}u$; as the s_i^σ are the same as the s_i except for a permutation, we can write them as $s_i^\sigma = s_{\tau(i)}$, where $i \rightarrow \tau(i)$ is a permutation. Then we have $\Phi(s_{\tau(i)}u^\sigma) = \Phi(s_{\sigma(i)}u)$, which can be written as $\bar{s}_{\tau(i)}\Phi(u^\sigma) = \bar{s}_{\sigma(i)}\bar{u}$, i. e. $\Phi(u^\sigma) = \bar{s}_{\tau(i)}^{-1}\bar{s}_{\sigma(i)}\bar{u}$. As $\boldsymbol{G}$ acts faithfully on $\boldsymbol{S}$, the corollary of Prop. 5 shows that all the elements $\bar{s}_{\tau(i)}^{-1}\bar{s}_{\sigma(i)}$ of $\boldsymbol{G}$, for $1 \leqq i \leqq r$, must coincide; if $\bar{t}$ is their common value, we have $\bar{s}_{\sigma(i)} = \bar{s}_{\tau(i)}\bar{t}$, which shows that the right-translation $\bar{t}$ maps the set $\bar{s}_i$ onto itself. By (S), this implies that $\bar{t}$ is the neutral element of $\boldsymbol{G}$, so that $\Phi(u^\sigma) = \bar{u}$, and therefore $u^\sigma = u$. As u is separably algebraic over $k'(y)$, this shows that $k'(u) \subset k'(y)$.

8. Proposition 7 shows that we may write $y = f(u)$, where f is a birational correspondence, defined over k', between W and the locus Y of y over k' in affine space. If $k'[y]$ is the ring generated over k' by the coordinates of y, it is well-known that the integral closure of $k'[y]$ in $k'(y)$ is a finitely generated ring over k', i. e. that it can be written as $k'[y^*]$, where y^* is a point in a suitable affine space; call Y^* the locus of y^* over k' in that affine space. As we have $k'(y^*) = k'(y) = k'(u)$, we may write $y^* = f^*(u)$, f^* being a birational correspondence between W and Y^*, defined over k'. It is usual to say that Y^* is derived from Y by "normalization" over k'. By Prop. 14 of the Appendix, since k'' is separably algebraic over k', $k''[y^*]$ is integrally closed in $k''(y^*)$.

PROPOSITION 8. *With the notations explained above, y^* and $\bar{u}$ are corresponding generic points over K'' on Y^* and $\boldsymbol{S}$ in a birational correspondence between Y^* and $\boldsymbol{S}$ which maps Y^* biregularly onto the K''-open set $\Omega = \bigcap_i \bar{s}_i^{-1}\Phi(W)$ on $\boldsymbol{S}$.*

In the first place, we prove that the coordinates $w_{i\mu}$ of the w_i are all in $k''[y^*]$; as they are in $k''(y)$ because of the relations $w_i = s_i u$ and $k'(u) = k'(y)$, it will be enough to show that they are integral over the ring $k''[y]$, or in other words (e.g. by F-App. II, Prop. 6) that they are everywhere finite on W. In fact, let π be any place of $k''(y)$ such that $y(\pi)$ is finite. Take r independent variables $\lambda_1, \cdots, \lambda_r$ over $k''(y)$, and extend π to a place π' of $k''(y, \lambda_1, \cdots, \lambda_r)$ at which every one of the r points $(\lambda_i, \lambda_i w_{i1}, \cdots, \lambda_i w_{im})$ is finite and $\neq (0, \cdots, 0)$. The relation (1), by which y was defined, can be written

$$\lambda_1 \cdots \lambda_r y(T, U) = \prod_{i=1}^{r} (\lambda_i T - \sum_{\mu} (\lambda_i w_{i\mu}) U_\mu).$$

Taking the values of both sides at π', we see that the right-hand side does not become identically 0 at that place; as $y(\pi)$ is finite, this implies that no λ_i can become 0 at π'; but then $w_{i\mu}(\pi)$ can be written as $(\lambda_i w_{i\mu})(\pi')/\lambda_i(\pi')$ and is finite. This proves the assertion about the $w_{i\mu}$.

We have thus shown that the mappings $y^* \rightarrow w_i$ of Y^* into W are everywhere defined on Y^*; as we have $\bar{u} = \bar{s}_i^{-1}\Phi(w_i)$, this implies that $y^* \rightarrow \bar{u}$ is everywhere defined and maps Y^* into the set Ω defined in Prop. 8. Conversely, the definition of y can be written

$$y(T, U) = \prod_{i=1}^{r} (T - \sum_{\mu} \Psi_\mu(\bar{s}_i \bar{u}) U_\mu)$$

if we call $\Psi_\mu(\bar{u})$ the coordinates of $\Psi(\bar{u})$. As Ψ is everywhere defined on $\Phi(W)$, this shows that the mapping $\bar{u} \rightarrow y$ is defined at every point of the set Ω. As $k'[y^*]$ is the integral closure of $k'[y]$ in $k'(y)$, it is therefore contained in the integral closure of the specialization-ring of every point of Ω on $\boldsymbol{S}$. But we have assumed that W and consequently $\boldsymbol{S}$ are normal, i.e. that the specialization-ring of every point of $\boldsymbol{S}$ (over any field of definition for $\boldsymbol{S}$) is integrally closed. This proves that $\bar{u} \rightarrow y^*$ is everywhere defined on the set Ω. In view of what we have proved above, Ω is therefore the set of points of $\boldsymbol{S}$ where this mapping is defined, and is K''-open by Prop. 8 of the Appendix; more precisely, it is K'-open if K' is the compositum of K and k'. This completes the proof.

9. Denote now by S any cycle $\sum_i (s_i)$ on V, *rational over the ground-field* k, consisting of distinct points s_i and satisfying condition (S). From such a set S, and taking $k' = k$, we can derive as above a point y, which we now write as y_S, and furthermore a point $y_S{}^*$ such that $k[y_S{}^*]$ is the integral closure of $k[y_S]$ in $k(y_S)$, as above, we call Y_S^* the locus of y_S^* over k; we write Ω_S for the open subset of $\mathbf{S}$ denoted by Ω in Prop. 8. If we allow S to run through any finite set of cycles with the properties stated above, then all the varieties Y_S^* will be birationally equivalent to W and to each other, and we can take the points y_S^* to be corresponding generic points of these varieties over k. It is then an immediate consequence of Prop. 8 that the affine varieties Y_S^* (with empty "frontiers"), and the birational correspondences between them for which the y_S^* are corresponding generic points of the Y_S^* over k, determine an abstract variety $\mathbf{S}'$, and that this is biregularly equivalent over a suitable field (as a matter of fact, over K itself) with the union of the open sets Ω_S on $\mathbf{S}$. In order to prove that $\mathbf{S}'$ will be biregularly equivalent to $\mathbf{S}$ itself for a suitable choice of the cycles S, it is therefore enough, in view of the well-known "compactoid" property of open sets in the Zariski topology, to show that the family of all open sets Ω_S *is* a covering of $\mathbf{S}$. In other words, we have to prove the following:

PROPOSITION 9. *Given any point $\bar{a}$ on $\mathbf{S}$, there is a cycle $S = \sum_i (s_i)$ on V, rational over k, consisting of distinct points s_i and satisfying condition* (S), *and such that $\bar{s}_i\bar{a} \in \Phi(W)$ for all i.*

Assume that $\bar{a}$ has a representative a_α on S_α; take x generic over $K(\bar{a}) = K(a_\alpha)$ on V, and put $u = (xt_\alpha^{-1})a_\alpha$, this being defined because W is a chunk. If we put, as usual, $\bar{x} = \Phi_0(x)$ and $\bar{u} = \Phi(u)$, we have then $\bar{u} = \bar{x}\bar{a}$, so that $u = \Psi(\bar{x}\bar{a})$. As the mapping $x \rightarrow \bar{x}\bar{a}$ is everywhere defined on V, this shows that the mapping $x \rightarrow u$ of V into W is defined at the points s of V such that $\bar{s}\bar{a} \in \Phi(W)$, and at those points only. Let F be the closed subset of V where the mapping $x \rightarrow u$ is not defined; by Prop. 12 of the Appendix, there is a maximal $\bar{k}$-closed subset F_0 of V contained in F; then an algebraic point of V over k is in F if and only if it is in F_0. Call F_1 the union of the conjugates over k of all the components of F_0; this is a k-closed set on V, and its definition shows that the cycle S on V will satisfy the last one of the conditions stated in Prop. 9 if and only if it lies in $V - F_1$.

Now assume first that the field k is infinite. Applying Prop. 13 of the Appendix to the variety $V' = V - F_1$, and to the empty subset of $V' \times V'$, we obtain a separably algebraic point s_1 over k on V'; call $s_1, \cdots, s_d$ all the distinct conjugates of s_1 over k; if this set satisfies condition (S), which will be the case in particular if $d = 1$, then it solves our problem. Suppose that this is not so, and therefore that $d > 1$. For any $r > d$, let $s_{d+1}, \cdots, s_r$ be any set of $r - d$ points on $V - F_1$, distinct from one another and from $s_1, \cdots, s_d$; put $S' = \{\bar{s}_1, \cdots, \bar{s}_d\}$ and $S'' = \{\bar{s}_{d+1}, \cdots, \bar{s}_r\}$. If the set

$S' \cup S''$ is mapped into itself by a right-translation τ other than the identity, one of the following circumstances must occur: (i) τ maps each one of the sets S', S'' onto itself; then τ is of the form $s'^{-1}t'$, with s', t' in S', and there must be two elements s'', t'' of S'' such that $t'' = s''\tau$; (ii) τ maps S' into S''; as $d > 1$, we can choose two distinct elements s', t' in S', and then $s'' = s'\tau$, $t'' = t'\tau$ are in S'', so that we have $t'' = (t's'^{-1})s''$; (iii) τ maps some $s' \in S'$ onto some $s'' \in S''$ and some $t' \in S'$ onto some $t_1' \in S'$; then $s'' = s't'^{-1}t_1'$. Thus, in order to satisfy the requirements of Prop. 9, it is enough to take as $s_{d+1}, \cdots, s_r$ the conjugates over k of a point $s = s_{d+1}$ of $V - F_1$, separably algebraic over k, satisfying the following conditions: (a) no $\bar{s}_l$, for $d + 1 \leqq l \leqq r$, coincides with any of the points $\bar{s}_i$ or $\bar{s}_i\bar{s}_j^{-1}\bar{s}_h$ for $1 \leqq i, j, h \leqq d$; (b) no pair of distinct conjugates of s over k lies on the graph of any of the birational correspondences $x \rightarrow \Psi(\bar{x}\bar{s}_i^{-1}\bar{s}_j)$, $x \rightarrow \Psi(\bar{s}_i\bar{s}_j^{-1}\bar{x})$ for $1 \leqq i, j \leqq d$. As to (a), it will be satisfied provided we take s on $V - F_2$, where F_2 is the union of F_1, of the set $s_1, \cdots, s_d$, and of the set of all conjugates over k of those algebraic points on V whose image on $\boldsymbol{G}$ coincides with one of the points $\bar{s}_i\bar{s}_j^{-1}\bar{s}_h$. Then our result follows at once by applying Prop. 13 of the Appendix to the variety $V - F_2$ and to the union of the graphs of the birational correspondences in (b).

If k is finite, we have to proceed differently. Take any algebraic point s_1 over k on $V - F_1$; call $s_1, \cdots, s_d$ its distinct conjugates over k; if this set satisfies condition (S), it solves our problem. If not, we use a result of Lang-Weil (this JOURNAL, vol. 76 (1954), p. 819) which says that, if l is sufficiently large, there must be a point s on $V - F_1$ which is rational over the (unique) extension of k of degree l. We take l prime and $> d$. If s is rational over k, the cycle (s) solves our problem; if not, it is of degree l over k; call $s_{d+1}, \cdots, s_{d+l}$ its distinct conjugates over k; they are distinct from $s_1, \cdots, s_d$, since the latter are of degree d over k. The set $s_{d+1}, \cdots, s_{d+l}$ may solve our problem. If it does not, the group $\mathfrak{g}$ of right-translations mapping the set $\{\bar{s}_{d+1}, \cdots, \bar{s}_{d+l}\}$ onto itself is of order $\nu > 1$; as that set must be the union of cosets with respect to $\mathfrak{g}$, ν must divide l, and so $\mathfrak{g}$ is cyclic of order l; call τ a generator of $\mathfrak{g}$. Let τ' be a right-translation mapping onto itself the set $\{\bar{s}_1, \cdots, \bar{s}_{d+l}\}$. If τ' is not the identity and maps some element of the set $\{\bar{s}_{d+1}, \cdots, \bar{s}_{d+l}\}$ into an element of the same set, it must be of the form τ^i, and therefore of order l; but this cannot be, since $d + l$ is not a multiple of l. Therefore τ' must map the set $\{\bar{s}_{d+1}, \cdots, \bar{s}_{d+l}\}$ into the set $\{\bar{s}_1, \cdots, \bar{s}_d\}$. As $d < l$, this is also impossible. Therefore the set $s_1, \cdots, s_{d+l}$ solves our problem.

This completes the proof of the results announced at the beginning of no. 7. Writing now G, S instead of $\boldsymbol{G'}$, $\boldsymbol{S'}$, we may restate them in a somewhat more complete form as follows:

THEOREM. (i) *To every pre-group V, defined over a field k, there is a birationally equivalent group G, also defined over k; this is uniquely determined up to an isomorphism.*

(ii) *To every pre-homogeneous space W with respect to V, defined over k, there is a birationally equivalent homogeneous space with respect to G, also defined over k; this is uniquely determined up to an isomorphism.*

(iii) *Let W be a pre-transformation space with respect to V, defined over k; let a be a point of W such that W is normal at a and that, if x is generic over $k(a)$ on V, xa and $x^{-1}(xa)$ are defined. Then there is a transformation-space S with respect to G, birationally equivalent to W over k in such a way that the birational correspondence between them is biregular at a; S may be taken everywhere normal, and it may be taken to be non-singular if a is simple on W. Moreover, S is uniquely determined up to a birational correspondence which is biregular at every point of the form $\bar{s}\bar{a}$, where $\bar{a}$ is the point corresponding to a on S and $\bar{s}$ is any point of G.*

Except for the statements about unicity, all this has been proved above. As to unicity, the statements in (i) and (ii) are special cases of the statement in (iii); and the latter is an immediate consequence of the fact that the operations of G are everywhere biregular mappings of S onto S.

Appendix.

If X is any cycle, we denote by $|X|$ the *support* of X, i. e. the closed set which is the set-theoretic union of the components of X.

PROPOSITION 1. *Let $k(x)$ be a regular extension of a field k, and $k(x,y)$ a regular extension of $k(x)$. Then $k(x,y)$ is a regular extension of k.*

This is an immediate consequence of F-I_7, Th. 5.

PROPOSITION 2. *Let $k(x)$ be a regular extension of a field k; let K be an overfield of k, linearly disjoint from $k(x)$ over k; let k' be the algebraic closure of k in K. Then $k'(x)$ is the algebraic closure of $k(x)$ in $K(x)$.*

Let y be an element of $K(x)$, algebraic over $k(x)$; we may take x to be a generic point over K of a variety V, defined over k, in an affine space; and then we may write $y = F(x)$, where F is a function on V, defined over K; call Γ the graph of F. As y is algebraic over $k(x)$, there is a polynomial $P \epsilon k[X, Y]$ such that $P(x, Y) \neq 0$ and $P(x, y) = 0$; then P induces on the product $V \times D$ of V and of the affine space D of dimension 1 a function which is not 0 on $V \times D$ and is 0 on Γ. As Γ has the same dimension as V, it must be a component of the divisor (P) of P, and is therefore algebraic over k. The smallest field of definition of Γ containing k must then be contained in k', so that F is defined over k'; this implies that y is in $k'(x)$.

COROLLARY. *If K is primary over k, $K(x)$ is primary over $k(x)$.*

In fact, the assumption means that k' is purely inseparable over k; this implies that $k'(x)$ is purely inseparable over $k(x)$.

PROPOSITION 3. *Let $k(x)$ be a finitely generated extension of a field k; then every field K such that $k \subset K \subset k(x)$ is finitely generated over k.*

Let $t = (t_1, \cdots, t_n)$ be a maximal set of algebraically independent elements of K over k; then K is algebraic over $k(t)$. Replacing k by $k(t)$, we see that it is enough to prove our proposition in the case when K is algebraic over k. This being assumed, call k' the smallest field of definition containing k for the locus of x over the algebraic closure $\bar{k}$ of k; then k' is a finite algebraic extension of k and is algebraically closed in $k'(x)$ since $k'(x)$ is regular over k'. But then k' is the algebraic closure of k in $k'(x)$ and therefore contains the algebraic closure of k in $k(x)$, so that K is contained in k'.

COROLLARY. *If $k(x)$ is regular over k, so is K.*

PROPOSITION 4. *Let t be a point, k a field, and let $t_1, \cdots, t_N$ be N independent generic specializations of t over k. Let x be a point of dimension $d < N$ over k and such that $k(x)$, $k(t)$ are linearly disjoint over k. Then there is an α such that t_α is a generic specialization of t over $k(x)$.*

Call n the dimension of $k(t)$ over k. By F-I$_6$, Th. 3, every t_α is a specialization of t over $k(x)$; if none is generic, every t_α must have over $k(x)$ a dimension $\leqq n-1$; but then $(x, t_1, \cdots, t_N)$ has over k a dimension $\leqq d + N(n-1) < Nn$, which is impossible, since $(t_1, \cdots, t_N)$ has the dimension Nn over k.

PROPOSITION 5. *Let V be a variety, defined over a field k; let K be an overfield of k and x a point of V. Let A and A' be the prime rational cycles,* *over k and over K respectively, with the generic point x. Then A is the same as A' if and only if K and $k(x)$ are linearly disjoint over k.*

We may replace V by any representative of V on which x has a representative, so that it is enough to prove our result for cycles in the affine n-space. For A to be the same as A', it is at any rate necessary that they should have the same dimension, so that K and $k(x)$ must be independent over k; assume from now on that this is so. Among the coordinates of x, let $(x_1, \cdots, x_r)$ be a maximal set of independent variables over k and therefore also over K; write y for the point $(x_1, \cdots, x_r)$ and z for $(x_{r+1}, \cdots, x_n)$. By F-VII$_6$, Th. 12, $A = A'$ if and only if $A \cdot (y \times S^{n-r})$ is the same as $A' \cdot (y \times S^{n-r})$; by F-VI$_3$, Th. 12, this is so if and only if z has the same complete set of conjugates over $K(y)$ as over $k(y)$, and therefore, by F-I$_4$, Prop. 12 and F-I$_2$, Prop. 6, if and only if $K(y)$ and $k(y, z) = k(x)$ are linearly disjoint over $k(y)$. The latter condition means that there is no relation $\sum_i u_i \Phi_i(y) = 0$ in which the u_i are linearly independent elements of $k(x)$ over $k(y)$ and the $\Phi_i(y)$ are in $K[y]$ and not all 0. Assume that there is such a relation; we may write $\Phi_i(y) = \sum_j P_{ij}(y)\xi_j$, where the ξ_j are linearly independent elements of K over k and the $P_{ij}(y)$ are in $k[y]$ and not all 0. Then we have $\sum_j v_j \xi_j = 0$ with $v_j = \sum_i u_i P_{ij}(y)$; as the v_j are in $k(x)$ and not all 0 because of the assumptions on the u_i and $P_{ij}(y)$, this shows that, when that is so, K and $k(x)$ are not linearly disjoint over k. Conversely, assume that there is a relation $\sum_j v_j \xi_j = 0$ in which the ξ_j are linearly independent elements of K over k and the v_j are in $k(x)$ and not all 0; as the ξ_j are then also linearly independent elements of $K(y)$ over $k(y)$, this implies that $K(y)$ and $k(x)$ are not linearly disjoint over $k(y)$.

COROLLARY. *Let V be a variety, defined over a field k. Let A be a prime rational cycle on V over an overfield K of k. Then, if K' is any field such that $k \subset K' \subset K$ over which A is rational, A is prime rational over K'; of all such fields K', there is one smallest one K_0; and an automorphism σ of K over k transforms A into itself if and only if it induces the identity on K_0.*

As in the proof of Prop. 5, it is enough to consider cycles in an affine space. Assume that A is prime rational over K and rational over $K' \subset K$, and write it as $A = \sum_i n_i A_i$, where the A_i are distinct prime rational cycles over K'. Let Z be a component of A_1; it is algebraic over K', and so every conjugate of Z over K is a fortiori such over K', so that every component of

A is a component of A_1; therefore we must have $A = n_1 A_1$. By F-I_8, Prop. 26, the coefficient of Z in A is at most equal to its coefficient in A_1; therefore we have $A = A_1$. That being so, it follows from Prop. 5 and from F-I_6, Th. 3 and F-I_7, Lemma 2, that there is a smallest field K_0 with the properties stated in our corollary; in fact, if x is a generic point of A over K, and if $\mathfrak{P}$ is the prime ideal in $K[X]$ consisting of all polynomials in $K[X]$ which are 0 at x, K_0 is the smallest subfield of K such that $\mathfrak{P}$ has a set of generators in $K_0[X]$. As $\mathfrak{P}$ is also the ideal in $K[X]$ whose set of zeros is the support $|A|$ of A, the last assertion follows from F-I_7, Lemma 2.

PROPOSITION 6. *Let V be a variety, defined over a field k, and A a cycle on V; assume either that A is a divisor on V or that the coefficients in A of all the components of A are $\not\equiv 0 \bmod p$, p being the characteristic. Then, of all the overfields of k over which A is rational, there is one smallest one k_0, k_0 is finitely generated over k; and an isomorphism σ of k_0 over k onto some extension of k leaves A invariant if and only if it leaves every element of k_0 invariant.*

Except for the last statement, this result is due to Chow. Let A be any cycle on V; for every representative V_α of V, call A_α the sum of the terms in the reduced expression for A which pertain to components with representatives in V_α; then A is rational over an overfield K of k if and only if every A_α is rational over K; and an isomorphism of K which leaves A invariant must leave all the A_α invariant. Therefore it is enough to deal with cycles on an affine variety V. For such a cycle A, put $A = \sum_n nA_n$, where A_n is the sum of the terms with the coefficient n in the reduced expression for A; then A is rational if and only if every cycle nA_n is rational; and an isomorphism which leaves A invariant must leave all the A_n invariant. Finally, if $n = p^\nu n'$ with n' prime to p, nA is rational if and only if $p^\nu A$ is rational. Therefore it will be enough to deal with the following two cases: (i) A is a cycle in affine space, consisting of a sum of distinct components; (ii) A is a divisor on an affine variety V and of the form $A = qA_0$, where q is a power of p and A_0 is a sum of distinct components.

(i) Let $\mathfrak{A}$ be the ideal of all polynomials (with coefficients in the universal domain) which are 0 on the support $|A|$ of A; this is the intersection of the prime ideals determined similarly by the components of A. The first assertion in our proposition will then be a consequence of F-I_7, Lemma 2, if we prove that A is rational over a field K if and only if $\mathfrak{A}$ has a set of generators in $K[X]$, i. e. if it is the extension to the universal domain of the ideal $\mathfrak{A} \cap K[X]$; the second assertion in our proposition also follows from the same lemma, provided one observes that, if k_0 is the smallest field such that $\mathfrak{A}$ has a set of generators in $k_0[X]$, an isomorphism which leaves A invariant must map k_0 onto k_0, i.e. it must induce an automorphism in k_0, so that the lemma in question is applicable.

If $\mathfrak{A}$ has a set of generators (P_ν) in $K[X]$, the support $|A|$ of A is the set of zeros of the P_ν and is therefore K-closed. On the other hand, $|A|$ must also be K-closed if A is rational over K. In order to prove the equivalence of those two properties, one may then begin by assuming that $|A|$ is K-closed. Consider first the case in which all the components of A are the conjugates of one of them, say Z, over K; let x be a generic point of Z over $\bar{K}$; then A is rational over K if and only if $K(x)$ is separable over K. Put $K' = K^{p^{-\infty}}$, this being the smallest "perfect" field containing K. Put:

$$\mathfrak{P} = \mathfrak{A} \cap K[X], \qquad \mathfrak{P}' = \mathfrak{A} \cap K'[X],$$

and call $\mathfrak{Q}'$ the extension of $\mathfrak{P}$ to $K'[X]$. By F-IV$_2$, Th. 4, and F-II$_1$, Prop. 3, $\mathfrak{P}$ and $\mathfrak{P}'$ consist of the polynomials, in $K[X]$ and in $K'[X]$ respectively, which are 0 at x; they are prime ideals; moreover, if $P' \in \mathfrak{P}'$, some power P'^n of P' is in $\mathfrak{P}'$ and hence in $\mathfrak{Q}'$; as $\mathfrak{Q}' \subset \mathfrak{P}'$, this implies that $\mathfrak{Q}'$ is primary and belongs to the prime ideal $\mathfrak{P}'$. By F-I$_6$, Th. 3, and F-I$_7$, Prop. 19, we see that $\mathfrak{P}' = \mathfrak{Q}'$ if and only if $K(x)$ is separable over K, and therefore, as we have shown, if and only if A is rational over K. But, if $\mathfrak{A}$ is the extension of $\mathfrak{P}$ to the universal domain, $\mathfrak{P}'$ must a fortiori be the extension of $\mathfrak{P}$ to $K'[X]$. Conversely, if $\mathfrak{Q}' = \mathfrak{P}'$, the extension of $\mathfrak{P}$ to the universal domain is the same as that of $\mathfrak{P}'$; but it is well-known and easily verified that the latter must be a "radical" ideal, i.e. one consisting of all the polynomials which are 0 on a closed set; then one sees at once that it must be the same as $\mathfrak{A}$. This completes the proof in the special case we were considering.

Now assume that $|A|$ is any K-closed set; then we can write A as the sum of cycles A_i such that the components of each A_i are mutually conjugate over K, and $\mathfrak{A}$ is the intersection of the ideals $\mathfrak{A}_i$ similarly determined by the A_i. Put:

$$\mathfrak{P}_i = \mathfrak{A}_i \cap K[X], \qquad \mathfrak{P}_i' = \mathfrak{A}_i \cap K'[X],$$

and call $\mathfrak{Q}_i'$ the extension of $\mathfrak{P}_i$ to $K'[X]$. If A is rational over K, all the A_i must be so, so that, as shown above, the $\mathfrak{A}_i$ must be the extensions of the $\mathfrak{P}_i$ to the universal domain. It is then easily seen that $\mathfrak{A}$ is the extension of the intersection of the $\mathfrak{P}_i$, i.e. of $\mathfrak{A} \cap K[X]$. Assume, on the other hand, that A is not rational over K; then we have $\mathfrak{Q}_i' \neq \mathfrak{P}_i'$ for at least one i; from the unicity of the decomposition of an ideal into an intersection of primary ideals, it follows then that the intersection of the $\mathfrak{Q}_i'$, which is the extension of $\mathfrak{A} \cap K[X]$ to $K'[X]$, cannot be the same as the intersection of the $\mathfrak{P}_i'$, which is $\mathfrak{A} \cap K'[X]$. A fortiori, $\mathfrak{A}$ cannot then be the extension of $\mathfrak{A} \cap K[X]$ to the universal domain. This completes the proof for case (i).

(ii) Let V be a variety, defined over k, in an affine space; let A_0 be a divisor on V and the sum of distinct components; let q be a power of the characteristic $p \neq 0$; put $A = qA_0$. If P is any polynomial which is not 0 on V, denote by $(P)_V$ the divisor of the function induced by P on V. Call $\mathfrak{A}$ the ideal of all the polynomials P, with coefficients in the universal domain, such that either $P = 0$ on V or $(P)_V \succ A$. If A is rational over an overfield K of k, $\mathfrak{A}$ is then the extension of $\mathfrak{A} \cap K[X]$ to the universal domain, as follows at once from F-VIII$_3$, Th. 10. Conversely, assume that $\mathfrak{A}$ is the extension of $\mathfrak{A} \cap K[X]$ to the universal domain; we will prove that A is then rational over K; our proposition will then follow from this as in case (i). As a polynomial P is 0 on $|A|$ if and only if some power P^n of P is in $\mathfrak{A}$, our assumption on A implies that A is K-closed, and therefore that A_0 is rational over $K' = K^{p^{-\infty}}$. Let Z be a component of A. As well-known, there is a polynomial P such that $(P)_V = A_0 + B$, where B has no component in common with A_0; write P as $P = \sum_i \xi_i P_i$, where the ξ_i are linearly independent over K' and the P_i are in $K'[X]$; by F-VIII$_3$, Th. 10, we have $(P_i)_V \succ A_0$ for all i; and Z must have the coefficient 1 in at least one of the P_i, since otherwise it would occur in B; if we call that polynomial P', P' is then in $K'[X]$, Z has the coefficient 1 in $(P')_V$, and we have $(P')_V \succ A_0$. But then P'^q is in $\mathfrak{A}$, and therefore, by hypothesis, may be written as $\sum_j \eta_j Q_j$, where the Q_j are in $\mathfrak{A} \cap K[X]$. The latter fact implies that Z has at least the coefficient q in all the $(Q_j)_V$; as it has the coefficient q in P'^q, it must have the coefficient q in one at least of the divisors $(Q_j)_V$; as these divisors are rational over K, this implies that, if A_1 is the sum of Z and its conjugates over K, qA_1 is rational over K. As this is so for every component Z of A, A is therefore rational over K.

PROPOSITION 7. *Let U, V be two varieties, defined over a field k; let F be a k-closed subset of $U \times V$. Then the set A of the points a on U such that $a \times V \subset F$ is k-closed.*

Let $W_1, \cdots, W_m$ be those components of F which have the "projection" V on V (in the sense of F-IV$_3$, F-VII$_3$); if v is a generic point of V over k, W_i has a generic point over $\bar{k}$ of the form (u_i, v); and $a \in A$ if and only if, for v' generic over $k(a)$ on V, (a, v') is a specialization of some (u_i, v) over $\bar{k}$. Let V_1 be any representative of the abstract variety V; let v_1 be the representative of v on V_1; the ambient affine space for V_1 being embedded in a projective space, let V_0 be the locus of v_1 over k in that projective space. Let F_0 be the union of the loci of the points (u_i, v_1) over $\bar{k}$ in $U \times V_0$; F_0 is k-closed on $U \times V_0$. Then A is the set of the points a on U such that $F_0 \cap (a \times V_0)$ has a component of dimension $\geqq \dim(V_0)$. As V_0 is complete, our conclusion is now contained in Lemma 7 of my paper in *Math. Ann.*, vol. 128 (1954), p. 104.

PROPOSITION 8. *Let ϕ be a mapping of a variety U into a variety V; let k be a field of definition for U, V and ϕ. Then the set of points of U where ϕ is defined is k-open.*

(i) Assume first that U is an affine variety and V is the affine space of dimension 1. Let x be a generic point of U over k; put $y = \phi(x)$. Let $\mathfrak{A}$ be the set of all polynomials P in $k[X]$ such that $P(x)y$ is in $k[x]$; this is an ideal in $k[X]$, containing the ideal $\mathfrak{P}$ of those polynomials which are 0 at x and therefore on V. Since y may be written as $Q(x)/P(x)$, with P, Q in $k[X]$ and $P(x) \neq 0$, we have $\mathfrak{A} \neq \mathfrak{P}$. As the points where ϕ is not defined are the zeros of $\mathfrak{A}$, the set of such points is k-closed.

(ii) Take V as in (i), and assume that U is an abstract variety, with the representatives U_α, on each of which a "frontier" F_α (i.e. a k-closed set) is given, according to the definitions in F-VII$_1$. Call F_α' the k-closed subset of U_α where ϕ is not defined; the set F of the points of U where ϕ is not defined is then the union of the images of the sets $F_\alpha' \cap (U_\alpha - F_\alpha)$ by the canonical birational mappings of the U_α into U. It is easily seen that F must be k-closed provided the following assertion is true: if x is a point of U with a representative x_α on some U_α which is a generic point over $\bar{k}$ of a component of F_α', then every specialization x' of x over k is in F. In fact, let β be such that x' has a representative x_β' on U_β; then x must also have a representative x_β on U_β, and, from the biregularity of the correspondence between U_α, U_β at (x_α, x_β), it follows that x_β must be in F_β'; as x_β' is a specialization of x_β over k, and as F_β' is k-closed, x_β' must then be in F_β', so that x' is in F. This proves our result for this case.

It follows trivially from this that our result remains true when U is an abstract variety and V is an affine space or more generally an affine variety.

(iii) Let U be an abstract variety and let V be a k-open subset of an affine variety V_1; let V_0 be the projective variety whose part "at finite distance" is V_1; then $V_0 - V$ is a k-closed subset F_0 of V_0. Call Γ the graph of ϕ on $U \times V_0$; the set F of the points of U where ϕ, considered as a mapping of U into V, is not defined, is then the union of the set F_1 of the points of U where ϕ is not defined as a mapping of U into V_1 and of the set-theoretic projection of $\Gamma \cap (U \times F_0)$ on U. As V_0 is complete, the latter set coincides with the "projection" in the sense of F-IV$_3$ and F-VII$_3$ and is k-closed (e. g. by F-VII$_4$, Prop. 10 and 11); and F_1 is k-closed, as shown in (ii). Therefore F is k-closed.

(iv) Let U, V be arbitrary abstract varieties; let x be a generic point of U over k; let the V_α be those representatives of V on which $\phi(x)$ has a representative $\phi_\alpha(x)$, and let F_α be the "frontiers" on the V_α. Then ϕ is defined at a point of U if there is an α such that ϕ_α, considered as a mapping of U into $V_\alpha - F_\alpha$, is defined there. Therefore the set where ϕ is not defined is the intersection of the sets where the ϕ_α are not defined; as the latter sets are k-closed by (iii), this completes the proof.

COROLLARY 1. *Let V be an abstract variety, defined over k, with the representatives V_α. Then, for each α, the set Ω_α of the points of V which have a representative on V_α is k-open; and the canonical correspondence between V and V_α is an everywhere biregular mapping of Ω_α onto $V_\alpha - F_\alpha$ if F_α is the frontier for V_α.*

Let x be a generic point of V over k, and let x_α be its representative on V_α; if we put $x_\alpha = \phi_\alpha(x)$, ϕ_α is the "canonical correspondence" between V and V_α. Then Ω_α is the set of points where ϕ_α, considered as a mapping of V into $V_\alpha - F_\alpha$, is defined; it is k-open by Prop. 8. The rest is obvious.

COROLLARY 2. *Let V be an abstract variety, defined over a field k. Then there is a finite covering of V by k-open subsets of V, each of which is biregularly equivalent to an affine variety.*

Corollary 1 says that V has a covering by the k-open sets Ω_α, each of which is biregularly equivalent to the k-open subset $V_\alpha - F_\alpha$ of the affine variety V_α. It is therefore enough to prove our assertion for a k-open subset $V - F$ of an affine variety V defined over k. Let $\mathfrak{A}$ be the set of all polynomials in $k[X]$ which are 0 on F; it is an ideal in $k[X]$, and, as F is k-closed, it is the set of zeros of $\mathfrak{A}$. Let $P_1, \cdots, P_m$ be a set of generators for $\mathfrak{A}$; as $F \neq V$, they are not all 0 on V, and we may assume that $P_1, \cdots, P_r$ are not 0 on V while $P_{r+1}, \cdots, P_m$ are 0 on V, with $1 \leqq r \leqq m$. For $1 \leqq \rho \leqq r$, call Ω_ρ the k-open subset of V consisting of the points where P_ρ is not 0; the Ω_ρ are a covering of $V - F$. Let $x = (x_1, \cdots, x_n)$ be a generic point of V over k; let V_ρ be the locus of the point

$$(x_1,\cdots,x_n,1/P_\rho(x_1,\cdots,x_n))$$

in the affine space of dimension $n+1$. Then V_ρ is biregularly equivalent to Ω_ρ.

COROLLARY 3. *Let V be a variety, defined over a field k. The set of points of V where V is normal* (resp. *relatively normal with respect to k*) *is a k-open subset of V.*

Let V^* be the variety derived from V by normalization with reference to the smallest perfect field $k'=k^{p^{-\infty}}$ containing k (resp. with reference to k); let x be a generic point of V over k; let x^* be the corresponding point of V^*, which is generic over k' (resp. over k) on V^*. We may then write $x^*=\phi(x)$, where ϕ is a mapping of V into V^*, defined over k' (resp. over k). Then the points where V is normal (resp. relatively normal) are those where ϕ is defined. As any k'-open set is also k-open, this proves the corollary.

PROPOSITION 9. *Let V be a variety, defined over a field k; let F be a closed subset of V. For F to be k-closed, it is necessary that it should contain all the specializations over k of all its points; it is sufficient that it should contain all the generic specializations over k of all its points, or also that it should be invariant under all isomorphisms over k of a common field of definition $K\supset k$ for its components.*

The necessity of the first condition follows from F-IV$_2$, Th. 4; we first prove that this condition is sufficient. In fact, it implies that, if z is a generic point over K of a component Z of F, the locus Z' of z over $\bar{k}$ is contained in F; as Z is the locus of z over $\bar{K}$, Z' contains Z; as z cannot be in any other component of F than Z, we get $Z'=Z$; thus all components of F are algebraic over k, and then F-IV$_2$, Th. 4, shows that all the conjugates of Z over k must be contained in F. Now we show that the second condition implies the first one. Let x be any point of F and let x' be a specialization of x over k. Then if V is the locus of x over $\bar{k}$, F-IV$_2$, Th. 4, shows that x' must be on a conjugate V' of V over k. Let x'' be a generic point of V' over $\bar{K}$; then x'' is a generic specialization of x over k by F-IV$_2$, Th. 4, and is therefore in F by hypothesis, and x' is a specialization of x'' over $\bar{K}$ and a fortiori over K and so is in F since F is K-closed. Finally the last condition implies the second one; for let x' be a generic specialization over k of a point x in a component Z of F; then the isomorphism of $k(x)$ onto $k(x')$ over k which maps x onto x' can be extended to an isomorphism σ of $K(x)$

onto a field $K^\sigma(x')$, and then x' is on Z^σ by F-IV$_2$, Th. 3, Coroll. 2, and is therefore in F if F is invariant under σ.

PROPOSITION 10. *Let W be a subvariety of a product $U \times V$, with the "projection" U on U; let k be a field of definition for U, V, W. Then the set-theoretic projection of W on U contains an open subset of U; and the union of all such sets is k-open.*

The assumption means that, if (u, v) is a generic point of W over k, u is generic over k on U. Let V_1 be a representative of V on which v has a representative v_1; F_1 being the corresponding frontier, put $V_1' = V_1 - F_1$, so that v_1 is in V_1'; let W_1 be the locus of (u, v_1) over k on $U \times V_1'$. Let V_0 be the projective variety whose part "at finite distance" is V_1; put $F_0 = V_0 - V_1'$; this is a k-closed set on V_0. The set $W_1 \cap (u \times V_1')$ can be written as $u \times X$, where X is either V_1' (in the trivial case $W = U \times V$) or else a $k(u)$-closed subset of V_1'; as v_1 is in X, X is not empty, so that we can choose in it a point w which is algebraic over $k(u)$. Let W' be the locus of (u, w) over $\bar{k}$ on $U \times V_0$, which has the same dimension as U; call n that dimension. Then the set $C = W' \cap (U \times F_0)$ is a $\bar{k}$-closed subset of W', so that all its components are of dimension $\leqq n - 1$. As V_0 is complete, the set-theoretic projection C' of C on U is then a $\bar{k}$-closed subset of U. Let a be any point in $U - C'$; as V_0 is complete, there is a point (a, b) on W' with the projection a on U; as a is not in C', b cannot be in F_0 and is therefore in V_1', so that (a, b) is in W_1. Therefore the $\bar{k}$-open set $U - C'$ on U is contained in the set-theoretic projection of W_1 and a fortiori in that of W. The last assertion in our proposition is then an immediate consequence of the sufficiency of the last condition in Prop. 9.

PROPOSITION 11. *Let U, V, W be three varieties and f a mapping of $U \times V$ into W, all defined over a field k. Assume that, for every $a \in U$, f is defined at (a, x) for x generic on V over $k(a)$. Let Ω be the set of those $a \in U$ such that, for x generic over $k(a)$ on V, $f(a, x)$ is generic over $k(a)$ on W. Then Ω is either empty or k-open on U.*

Let r be the dimension of W; for z generic over k on W, let $z_1, \cdots, z_r$ be r algebraically independent elements of $k(z)$ over k; put $z_i = \phi_i(z)$, where ϕ_i is a function on W, defined over k. It is clear that a point z' of W is generic over an overfield K of k if and only if the ϕ_i are all defined at z' and their values $\phi_i(z')$ are independent over K. Let u, x be independent generic points of U, V over k; we may assume that $f(u, x)$ is generic over k on W, since otherwise Ω is empty. Put $f_i = \phi_i \circ f$; Ω is then the set of those points a on U such that, for x generic over $k(a)$ on V, the $f_i(a,x)$ are all defined and are algebraically independent over $k(a)$.

Take u, x as above; assume that u is not in Ω; we prove that Ω must then be empty. In fact, the assumption on u means that there is a polynomial P with coefficients in $k(u)$ such that

$$P(f_1(u,x),\cdots,f_r(u,x))=0.$$

Write $P=\sum_\nu t_\nu M_\nu(Z)$, where the $M_\nu(Z)$ are monomials (with coefficient 1) in the indeterminates $Z_1,\cdots,Z_r$, and the t_ν are in $k(u)$ and not all 0. Let a be any point of U, and take x' generic over $K=k(a)$ on V. Take a variable quantity λ over $k(u,x)$; extend the specialization $u\rightarrow a$ over k to a $\bar{K}$-valued place π of the field $k(u,\lambda)$ such that the elements λt_ν of $k(u,\lambda)$ are all finite and not all 0 at π; call t_ν' the value of λt_ν at π. As $k(u,\lambda)$ and $k(x)$ are independent regular extensions of k, the place π of $k(u,\lambda)$ and the isomorphism of $k(x)$ onto $k(x')$ over k which maps x onto x' make up a specialization of the set of quantities $k(u,\lambda)\cup k(x)$, which can be extended to a place π' of $k(u,\lambda,x)$ at which u, x and the λt_ν have respectively the values a, x' and t_ν'. If the $f_i(a,x')$ are not all defined, a is not in Ω; if they are all defined, they are the values at π' of the elements $f_i(u,x)$ of $k(u,\lambda,x)$. In the latter case, the relation

$$\lambda P(f_1(u,x),\cdots,f_r(u,x))=0,$$

taken at π', gives an algebraic relation between the $f_i(a,x')$ whose coefficients t_ν' are in $\bar{K}$ and are not all 0; this implies that the $f_i(a,x')$ are not independent over $K=k(a)$, so that a is again not in Ω. This shows that, for $u\notin\Omega$, Ω must be empty. From now on, therefore, we may assume that u is in Ω.

We prove now that Ω must contain a k-open set. Since the assumptions and the conclusion of our proposition are not affected if V is replaced by any birationally equivalent variety to V over k (the mapping f being transferred to the latter in an obvious manner), we may take for V an affine variety; put $x=(x_1,\cdots,x_m)$. Then we can write the f_i as

$$f_i(u,x)=P_i(x)/P_0(x),$$

where $P_0,P_1,\cdots,P_r$ are polynomials in the indeterminates $X_1,\cdots,X_m$ with coefficients in $k(u)$, and $P_i(x)\neq 0$ for $0\leqq i\leqq r$. Call $M_\nu(X)$, with $0\leqq\nu\leqq N$, all the monomials in $X_1,\cdots,X_r$ which either are of degree 0 or 1 (i. e. equal to one of the monomials 1, $X_1,\cdots,X_r$) or occur in one at least of the P_i; call $\bar{x}$ the point in the projective space P^N with the homogeneous coordinates $(M_\nu(x))$, and call $\bar{V}$ the locus of $\bar{x}$ over k. We may replace V by the birationally equivalent $\bar{V}$; then, writing V, x instead of $\bar{V}$, $\bar{x}$, and calling $(x_0, \cdots, x_N)$ the homogeneous coordinates for x, we see that the $f_i(u,x)$ are expressed as z_i/z_0, with

$$z_i = \sum_{\nu=0}^{N} t_{i\nu} x_\nu \qquad (0 \leqq i \leqq r),$$

where the $t_{i\nu}$ are elements of $k(u)$. If V is contained in any linear subvariety of P^N, then the smallest linear subvariety of P^N which contains V is defined over k; if this is of dimension N', we can express $N - N'$ of the coordinates x_ν linearly in terms of the others, with coefficients in k; thus we may assume that V is not contained in any linear subvariety of P^N.

We may write $t_{i\nu} = \phi_{i\nu}(u)$, where the $\phi_{i\nu}$ are functions on U, defined over k; as z_0 is not 0, we may assume that $t_{00} = 1$. By Prop. 8, the subset U' of U where all the $\phi_{i\nu}$ are defined and finite is k-open. Call n the dimension of V; as n is then the dimension of $k(u,x)$ over $k(u)$, and the $f_i(u,x)$ are independent over $k(u)$, we have $r \leqq n$. Put $z_j = \sum_\nu w_{j\nu} x_\nu$ for $r+1 \leqq j \leqq n$, where the $w_{j\nu}$, for $r+1 \leqq j \leqq n$, $0 \leqq \nu \leqq N$, are $(n-r)(N+1)$ independent variables over $k(u,x)$; call S the affine space of dimension $(n-r)(N+1)$. By F-II$_5$, Prop. 24, the $n-r$ quantities $z_{r+1}/z_0, \cdots, z_n/z_0$ are algebraically independent over the field

$$K = k(u, w, z_1/z_0, \cdots, z_r/z_0).$$

Now take any $a \in U'$; take $\bar{x}$, $\bar{w}$ generic and independent over $k(a)$ on V, S; put $\bar{t}_{i\nu} = \phi_{i\nu}(a)$, $\bar{z}_i = \sum_\nu \bar{t}_{i\nu}\bar{x}_\nu$ for $0 \leqq i \leqq r$, and $\bar{z}_j = \sum_\nu \bar{w}_{j\nu}\bar{x}_\nu$ for $r+1 \leqq j \leqq n$. As $\bar{t}_{00} = 1$, and V is not contained in any linear variety, $\bar{z}_0$ is not 0; therefore, if we put $f_j = z_j/z_0$ for $r+1 \leqq j \leqq n$, the functions $f_1, \cdots, f_n$ on $U' \times S$ are defined at $(a, \bar{w})$ and have the values $\bar{z}_1/\bar{z}_0, \cdots, \bar{z}_n/\bar{z}_0$ respectively. If one assumes that $\bar{z}_1/\bar{z}_0, \cdots, \bar{z}_n/\bar{z}_0$ are algebraically independent over $k(a, \bar{w})$ this implies a fortiori that $\bar{z}_1/\bar{z}_0, \cdots, \bar{z}_r/\bar{z}_0$ are so over $k(a)$, i.e. that $a \in \Omega$. Therefore, if we prove that there is a k-closed subset C of $U' \times S$ such that, with the notations just introduced, the quantities $\bar{z}_1/\bar{z}_0, \cdots, \bar{z}_n/\bar{z}_0$ are algebraically independent over $k(a, \bar{w})$ whenever $(a, \bar{w})$ is not in C, it will follow that Ω contains the set of those points a on U' such that $a \times S$ is not contained in C; and this set will be k-open by Prop. 7. In other words, as long as we merely wish to prove that Ω contains a k-open set, it is enough to prove it for $U' \times S$ and the functions $f_i = z_i/z_0$ for $1 \leqq i \leqq n$ instead of for U and for $f_1, \cdots, f_r$. This means that, writing U instead of $U' \times S$, it is enough to prove our assertion under the additional assumption $r=n$, the $\phi_{i\nu}$ being now everywhere defined on U, with $\phi_{00}=1$.

This being now assumed, put $z_{n+1}=\sum_\nu w_\nu x_\nu$, where the w_ν are $N+1$ independent variables over $k(u,x)$. As $k(u,w,x)$ is of dimension n over $k(u,w)$, there is a homogeneous polynomial P, with coefficients in $k(u,w)$ and not all 0, such that $P(z_0,\cdots,z_n,z_{n+1})=0$, and P is uniquely determined up to a factor in $k(u,w)$. As $z_1/z_0,\cdots,z_n/z_0$ are algebraically independent, there is at least one term in P where z_{n+1} occurs with a non-zero exponent; after multiplying P with a suitable element of $k(u,w)$, we may assume that the coefficient of this term is 1. Write all the other coefficients in P as $\psi_\rho(u,w)$, where the ψ_ρ are functions on $U\times S^{n+1}$, defined over k. We now prove our assertion about Ω by showing that Ω contains the set of all points a on U such that all the ψ_ρ are defined at $(a,\bar{w})$ for $\bar{w}$ generic on S^{n+1} over $k(a)$; this is a k-open subset of U by Prop. 8 and 7. In fact, let a be a point with that property; take $\bar{x}$ generic over $k(a,\bar{w})$ on V. Put $\bar{t}_{i\nu}=\phi_{i\nu}(a)$, these being all defined, according to our present assumptions; put $\bar{z}_i=\sum_\nu \bar{t}_{i\nu}\bar{x}_\nu$ for $0\leqq i\leqq n$, and $\bar{z}_{n+1}=\sum_\nu \bar{w}_\nu\bar{x}_\nu$. If we specialize the relation

$$P(z_0,\cdots,z_{n+1})=0$$

over the specialization $(\bar{a},\bar{w},\bar{x})$ of (u,w,x) with respect to k, we get, since the ψ_ρ are all defined at $(a,\bar{w})$, a homogeneous relation between $\bar{z}_0,\cdots,\bar{z}_{n+1}$ with coefficients in $k(a,\bar{w})$, containing $\bar{z}_{n+1}$ with a non-zero exponent in a term of coefficient 1. This shows that $\bar{z}_{n+1}/\bar{z}_0$ is then algebraic over the field $L(\bar{w})$, where we have put

$$L=k(a,\bar{z}_1/\bar{z}_0,\cdots,\bar{z}_n/\bar{z}_0).$$

Now take $n(N+1)$ independent variables $w_{i\nu}$ over $k(a)$, for $1\leqq i\leqq n$, $0\leqq\nu\leqq N$; put $y_i=\sum_\nu w_{i\nu}\bar{x}_\nu$ for $1\leqq i\leqq n$; what we have proved above shows that, for each i, $y_i/\bar{z}_0$ is algebraic over $L(w_{i0},\cdots,w_{iN})$, and therefore a fortiori over the field

$$L'=L(w_{10},\cdots,w_{nN})=k(a,w_{10},\cdots,w_{nN},\bar{z}_1/\bar{z}_0,\cdots,\bar{z}_n/\bar{z}_0).$$

On the other hand, one sees just as before, using F-II$_5$, Prop. 24, that the $y_i/\bar{z}_0$, for $1\leqq i\leqq n$, are algebraically independent over the field $k(a,w_{10},\cdots,w_{nN})$; as they are algebraic over L', this implies that L' has at least the dimension n over the latter field, so that the $\bar{z}_i/\bar{z}_0$, for $1\leqq i\leqq n$, must be algebraically independent over it. But then they must a fortiori be so over $k(a)$, which means that a is in Ω.

This completes the proof of the following statement: the assumptions being again those of Prop. 11, Ω must either be empty or contain a k-open subset of U. Now we prove Prop. 11 by induction on the dimension of U, the conclusion being trivially true if that dimension is 0. Assume that Ω is not empty; put $X = U - \Omega$; we have proved that X is contained in a k-closed subset C of U. Call U_i the components of C; they are algebraic over k, and of dimension $< \dim(U)$. By the induction assumption, $\Omega \cap U_i$ is either empty or a $\bar{k}$-open subset of U_i; in both cases its complement C_i on U_i is a $\bar{k}$-closed subset of U. As X is the union of the C_i, this shows that X is $\bar{k}$-closed. As it is obviously invariant by all automorphisms of $\bar{k}$ over k, it must then be k-closed.

PROPOSITION 12. *Let U be a variety defined over a field k; let F be a closed subset of U. Then, among all the k-closed subsets of U contained in F, there is one maximal set F_0.*

Let K be the smallest common field of definition containing k for all the components of F; let σ run through all the isomorphisms of K over k into the universal domain. As such an isomorphism σ leaves all k-closed sets invariant, every k-closed subset of U which is contained in F is contained in all the sets F^σ and therefore in their intersection F_0; F_0 is closed, since it is the intersection of closed sets; and it is k-closed, by Prop. 9. This proves the proposition.

PROPOSITION 13. *Let U be a variety defined over an infinite field k; let F be a closed subset of $U \times U$. Then there is a point a on U, separably algebraic over k and such that no pair (a', a'') of distinct conjugates of a over k is in F.*

Applying Prop. 12 to $U \times U$, F and the algebraic closure $\bar{k}$ of k, we see that there is a $\bar{k}$-closed subset F_0 of $U \times U$ such that a subvariety of $U \times U$ which is algebraic over k is contained in F if and only if it is contained in F_0; this applies in particular to algebraic points over k on $U \times U$. By replacing F by the union of all conjugates over k of all the components of F_0, we see that it is enough to prove our result in the case in which F is k-closed. We may assume that no component of F is contained in the diagonal of $U \times U$, since the omission of such components does not affect the content of our proposition. Furthermore, we may, in order to prove our proposition, replace U by any k-open subset of U; in view of this, we first replace U by the set of its simple points, and then use Corollary 2 of Prop. 8 to replace U by an affine variety. Thus we may assume that U is a non-singular affine variety, defined over k, and that F is a k-closed subset of $U \times U$, no component of which is contained in the diagonal of $U \times U$.

Let n and N be the dimensions of U and of the ambient affine space, respectively. The case $n = N$ is trivial, since in that case any rational point of U over k, e.g. 0, would solve our problem; therefore we assume $n < N$. Consider all sets of n linear equations:

$$(1) \qquad \sum_{\nu=1}^{N} t_{i\nu} X_\nu = t_{i0} \qquad (1 \leqq i \leqq n),$$

and identify the set (1) with the point $t = (t_{i0}, t_{i\nu})$ in the affine space T of dimension $n(N+1)$. In the space T, we consider the following sets:

(a) Call A the set of those points t for which the left-hand sides of (1) are not linearly independent; as A can be described as the set of zeros of certain determinants, it is k_0-closed, k_0 being the prime field (one could easily see that A is in fact a variety, defined over k_0). Put $T' = T - A$; for $t \in T'$, (1) defines a linear variety $L(t)$ of dimension $N - n$.

(b) Take t generic over k on T; by F-V_1, Th. 1, $U \cap L(t)$ is not empty, and, if u is a point in it, u is algebraic over $k(t)$ and is generic on U over the field $K = k(t_{11}, \cdots, t_{nN})$. As the t_{i0} are then in $K(u)$, we have $k(u,t) = K(u)$, so that $k(u,t)$ is a regular extension of k. Let W be the locus of (u,t) over k on $U \times T$; by F-V$_1$, Prop. 4, if $t' \in T'$, a point u' is in $U \cap L(t')$ if and only if (u',t') is in W. By Prop. 10, there is a k-closed subset B of T such that $T - B$ is contained in the set-theoretic projection of W on T; then, if $t' \in T - (A \cup B)$, $U \cap L(t')$ is not empty.

(c) Let $P_\rho(X) = 0$, for $1 \leqq \rho \leqq r$, be a set of equations for U with coefficients in k; put $P_{\rho\nu} = \partial P_\rho / \partial X_\nu$. Let D be the subset of $U \times T$ consisting of the points where the matrix

$$\left\| \begin{matrix} t_{i\nu} \\ P_{\rho\nu}(u) \end{matrix} \right\| \qquad (1 \leqq i \leqq n, 1 \leqq \rho \leqq r; 1 \leqq \nu \leqq N)$$

is of rank $< N$; since this can be expressed by the vanishing of determinants, D is a k-closed subset of $U \times T$ (as U is non-singular, it could be shown that D is actually a variety, defined over k). As W is not contained in D, $D \cap W$ is a k-closed subset of W (also, in fact, a variety), so that its components have a dimension $< n(N+1)$. Let D' be the "projection" of $D \cap W$ on T (i.e. the closure of the set-theoretic projection); this is a k-closed subset of T. Let u' be a point in $U \cap L(t')$, for $t' \in T'$; then, if $L(t')$ is not transversal to U at u', (u', t') must be in D and therefore in

$D \cap W$, and t' must be in D'. Therefore, if $t' \epsilon T - (A \cup D')$, $L(t')$ is transversal to U at every point of $U \cap L(t')$.

(d) Let X be any component of F; let (u, v) be a generic point of X over $\bar{k}$. As X is not in the diagonal of $U \times U$, we have $u \neq v$ and may assume for instance that $u_1 \neq v_1$. Take the $t_{i\nu}$ independent over $k(u, v)$ for $1 \leqq i \leqq n$, $2 \leqq \nu \leqq N$; as $n < N$, t will then be in T' for all choices of the t_{i1}, t_{i0}. Determine the t_{i1}, t_{i0} by the condition that $L(t)$ should contain both u and v; this determines them uniquely. As X is at most of dimension $2n - 1$, (u, v, t) is then of dimension $< n(N + 1)$ over k; therefore the locus Y of t over $\bar{k}$ is not T. For any $t' \epsilon T'$, assume that $U \cap L(t')$ contains two distinct points u', v' such that (u', v') is in X; then there is ν such that $u_\nu' \neq v_\nu'$, which implies that $u_\nu \neq v_\nu$. It is easily seen that the $t_{i\mu}$, for $1 \leqq i \leqq n$ and all $\mu \neq \nu$ and $\neq 0$, must then be independent over $k(u, v)$, and furthermore that (u', v', t') must be a specialization of (u, v, t) over $\bar{k}$, so that t' is in Y. Therefore, if t' is in T' and is not in the union C of all the varieties Y corresponding in this manner to the components X of F, there cannot be a pair of distinct points (u', v') in $U \cap L(t')$ such that (u', v') is in F. To conjugate components X, X' of F over k, there correspond conjugate varieties Y, Y' over k in T; therefore C is a k-closed subset of T.

Now let $P(t)$ be any polynomial other than 0 in the coordinates of t, with coefficients in k, which is 0 on the union of the k-closed subsets A, B, D' and C of T. As k is infinite, there is on T a rational point t over k such that $P(t) \neq 0$. As t is not in A, it determines a linear variety $L(t)$; as t is not in B, $U \cap L(t)$ is not empty. Take for a any point in $U \cap L(t)$; as t is not in D', $L(t)$ is transversal to U at a, so that a is separably algebraic over k. As all the conjugates of a over k are in $U \cap L(t)$, and as t is not in C, no pair of distinct conjugates a', a'' of a over k can be such that (a', a'') is in F. This completes the proof.

If k is a finite field, the conclusion of Prop. 13 need not be true. In fact, take for U a variety without rational points over k (e.g. the plane non-singular curve $x^4 + y^4 + z^4 = 0$ over the field with 5 elements); q being the number of elements of k, and $(x_1, \cdots, x_n)$ being a representative of a generic point of U over k, call x' the point whose corresponding representative is $(x_1^q, \cdots, x_n^q)$. Then the conclusion of Prop. 13 is false if we take for F the locus of (x, x') over k on $U \times U$.

PROPOSITION 14. *Let V be an affine variety, defined over a field k; let x be a generic point of V over k; assume that the ring $k[x]$ is integrally closed in $k(x)$. Then, if k' is any separably algebraic extension of k, $k'[x]$ is integrally closed in $k'(x)$.*

Put $n = [k' : k]$; as k' is separably algebraic over k, there are n distinct isomorphisms σ of k' into the algebraic closure of k. Each σ can then be extended uniquely to an isomorphism, which we also denote by σ, of $k'(x)$ onto $k'^{\sigma}(x)$ over $k(x)$. Let z be an element of $k'(x)$, integral over $k'[x]$ and therefore also over $k[x]$; then all the z^{σ} are also integral over $k[x]$. If $\xi_1, \cdots, \xi_n$ are n linearly independent elements of k' over k, it is well-known that $\det(\xi_i^{\sigma})$ is not 0; therefore all the z^{σ}, and z among them, can be expressed as linear combinations of the n elements $w_i = \sum_{\sigma} \xi_i^{\sigma} z^{\sigma}$; but these are integral over $k[x]$ and are traces over $k(x)$ of elements of $k'(x)$, so that they are in $k(x)$; as $k[x]$ is integrally closed, the w_i are therefore in $k[x]$, so that z is in $k'[x]$.

THE UNIVERSITY OF CHICAGO.

CORRECTIONS TO THE PAPER "LINEAR GRAPHS OF DEGREE $\leqq 6$ AND THEIR GROUPS."*

By I. N. Kagno.

D. W. Crowe and Frank Harary in a paper "Linear graphs through seven points . . ." which they are preparing for publication, call attention to a graph which was omitted by the author in his paper mentioned in the above title (this Journal, vol. 68, 1946, pp. 505-520, and vol. 69, 1947, p. 872). The author finds that this was inadvertently omitted in preparing his paper for publication, and now wishes to call attention to the following corrections to cover this omission.

On page 514, after line 16 add:

$$H'_7 \equiv [ab, ac, ad, ae, af, bc, bd, be, bf, ce, cf].$$

On page 515, after Theorem 3.6 add:

Theorem 3.6*. *H'_7 has the group* $(abc)\mathrm{all}(def)\mathrm{all}$.

On page 520, Theorem D.11 should read:

$$\mathfrak{H}_{28} \textit{ has the graphs } P \equiv [\cdot\ \cdot\ \cdot] \textit{ and } H'_7.$$

New York City.

---

* Received February 21, 1955.

ON STRONG SUMMABILITY OF FOURIER SERIES.*

By R. Salem.

Introduction. Let $f(x) \varepsilon L$ have period 2π and denote by $s_n(x)$ the partial sum of order n of the Fourier series of $f(x)$. Write

$$\phi_x(t) = f(x+t) + f(x-t) - 2f(x).$$

It is now a classical result that, if $f \varepsilon L^r$ $(r > 1)$, then at every point x where

$$\int_0^h |\phi_x(t)|^r \, dt = o(h), \tag{1}$$

(relation which is true for almost all x), one has

$$\sum_{\nu=0}^{n} |s_\nu(x) - f(x)|^q/(n+1) \to 0 \tag{2}$$

for all positive exponents q. See Zygmund [3] and the references quoted there. This theorem has been generalized by Marcinkiewicz [1] and by Zygmund [4] to the case in which one supposes only that $f \varepsilon L$.

In the case $q = 1$, the result can be considered as an extension of the classical Fejér summability theorem, to which it reduces if instead of the average of the absolute values $\sum_0^n |s_\nu - f|/(n+1)$, we consider the average without absolute values $\sum_0^n (s_n - f)/(n+1) = \sigma_n - f$ where σ_n denotes the Fejér sum of order n. Hence the name of "strong summability."

It is natural to ask the question whether the average (2) tends to zero, if instead of considering *all* sums $s_\nu(x)$ $(\nu = 0, 1, 2, \cdots)$, we consider only indices forming a subsequence of the sequence of natural numbers.

The first result in this direction is due to Zalcwasser [2] who proved that, if $f \varepsilon L$,

$$\sum_{k=0}^{n} (s_{k^2} - f)/(n+1) \to 0$$

---

* Received December 27, 1954.

at every Lebesgue point, and in particular at every point of continuity. This result was later extended to indices k^3 or k^4. All proofs are based on number theory methods using the arithmetical properties of the particular kind of sequence involved. Besides, the averages are not taken in absolute value.

In a different direction, some results where the averages are taken with absolute values have been obtained for more general sequences of indices by Zalcwasser and Zygmund but they are only known to hold almost everywhere (even for continuous f), the exceptional set of measure zero depending on the particular sequence $\{n_k\}$ of indices considered.

The main purpose of this paper is to show (in Theorem I) that if $f \varepsilon L^r$ $(r > 1)$ and if the condition (1) is satisfied at a point x, then at this point $\sum_{j=1}^{k} | s_{n_j}(x) - f(x) |^q / k \to 0$ for an arbitrarily large exponent q, and for a rather general kind of sequence $\{n_k\}$ which does not increase too rapidly and possesses some weak kind of regularity. Arithmetical properties of the sequence do not enter in the argument.

In the remarks following the proof of Theorem I we show that if $\{n_k\}$ increases rapidly enough, the theorem becomes false. We also show that some hypothesis about the regularity of the sequence can not be avoided.

Theorem II, although not directly conneceted with the previous results, is somewhat related to them; its object is to show that, denoting by M_n the maximum of $| s_n(x) |$ with respect to x, there exist continuous functions for which $(M_1 + \cdot\cdot\cdot M_n)/n$ is unbounded.

THEOREM I. *Let $\{n_k\}$ be an increasing sequence of positive integers satisfying the following conditions:*

1) *There exists a constant A such that $n_k = O(k^A)$*

2) *There exists a constant $\epsilon > 0$ such that $(n_{k+1} - n_k)/n_k > \epsilon/k$*

3) *There exists a lacunary sequence k_ν $(k_{\nu+1}/k_\nu \geqq R > 1)$ such that $n_{k_{\nu+1}}/n_{k_\nu} < B$, B being a constant.*[1]

Let $f(x)$ with period 2π belong to L^s $(s > 1)$, and denote by $s_m(x)$ the partial sum of order m of its Fourier Series. Write

$$\phi_x(t) = f(x+t) + f(x-t) - 2f(x).$$

---

[1] Condition 2) is certainly verified if $\{n_k\}$ is convex, and condition 3), if for a positive c, n_k/k^c decreases.

Then at every point x where $\int_0^\eta |\phi_x(t)|^s\,dt = o(\eta)$ *one has, as* $k\to\infty$,

$$\{|s_{n_1}(x)-f(x)|^r+\cdots+|s_{n_k}(x)-f(x)|^r\}/k\to 0$$

the exponent r being fixed, but arbitrarily large.

Preliminary remarks. a) It is well known that if the theorem is true for the exponent r, it is also true for any smaller exponent. Let q be the exponent complementary to r $(1/q+1/r=1)$. Taking r large enough we have $q<s$, so that we can assume $r\geqq 2$, $1<q\leqq 2$, $f\,\varepsilon\,L^q$ and

$$\int_0^\eta |\phi_x(t)|^q\,dt = o(\eta).$$

b) In the hypothesis 3) about the sequence $\{n_k\}$ we shall assume that $k_\nu = 2^\nu$; the changes in the proof for the general case are trivial.

c) We remark that

$$\int_0^\eta |\phi_x(t)|^s\,dt = o(\eta) \text{ implies } \int_0^\eta |\phi_x(t)|\,dt = o(\eta),$$

hence by a classical result

$$s_n - f = \frac{1}{\pi}\int_\delta^\pi \phi_x(t)\sin nt\,dt/t + o(1)$$

as $n\to\infty$, provided that $\delta = O(1/n)$.

Proof of the theorem. We shall assume temporarily that in the hypothesis 2), $\epsilon>2$, and we shall lift this restriction at the end of the proof.

Consider now an $h=2^\nu$. We shall prove first that $\sum_{j=h}^{j=2h-1}|s_{n_j}-f|^r = o(h)$ as $h\to\infty$. From this the theorem will follow easily.

By remark (c), and since $n_{2h}/n_h < B$ it will be enough to prove that $T=\sum_{j=h}^{2h-1}\left|\int_{\pi/n_h}^\pi \phi_x(t)\sin n_j t\,dt/t\right|^r = o(h)$. For this purpose we shall consider the sum

$$S=\sum_{j=h}^{2h-1}\alpha_j\int_{\pi/n_h}^\pi \phi_x(t)\sin n_j t\,dt/t = \int_{\pi/n_h}^\pi \phi(t)\,(\sum_h^{2h-1}\alpha_j\sin n_j t)\,dt/t,$$

where the coefficients α_j are arbitrary, and where we write $\phi_x(t)=\phi(t)$ since x is fixed and no ambiguity can arise.

We can write $n_h = h^{P+\rho}$ where $P=P(h)$ is an integer and $\rho=\rho(h)$ is such that $0<\rho\leqq 1$. Obviously $P\geqq 0$; also $P\leqq A$. But $P=0$ only if $\rho=1$.

The integrand being the same as in S, we write, if $P > 0$,

$$S = \int_{\pi/n_h}^{h\pi/n_h} + \cdots + \int_{\pi h^{p-1}/n_h}^{\pi h^p/n_h} + \cdots + \int_{\pi h^{P-1}/n_h}^{\pi h^P/n_h} + \int_{\pi/h^P}^{\pi}$$

$= U_1 + \cdots + U_p + \cdots + U_P + V$, say; if $P = 0$ we have merely $S = V$, with $\rho = 1$.

By a change of variable,

$$U_p = \int_{\pi/h}^{\pi} \phi(h^p\theta/n_h) \, (\sum_{j=h}^{2h-1} \alpha_j \sin(n_j h^p/n_h)\theta) \, d\theta/\theta.$$

We shall find a convenient bound for $|U_p|$. We remark that

$$(n_{j+1} - n_j) h^p/n_h \geqq (n_{j+1} - n_j) h/n_j > (n_{j+1} - n_j) j/2n_j > \epsilon/2 > 1$$

so that all integers $[n_j h^p/n_h]$ are different (we denote as usual by $[z]$ the integral part of z). To simplify the writing we shall study the integral

$$U = \int_{\pi/h}^{\pi} \phi(\lambda\theta) \, (\sum_{h}^{2h-1} \alpha_j \sin n_j\lambda\theta) \, d\theta/\theta,$$

where $0 < \lambda < 1$, $(n_{j+1} - n_j)\lambda > 1$, and $n_j\lambda \geqq h$. (All three conditions are satisfied by $\lambda = h^p/n_h$.)

We write $\delta_j = (n_j\lambda - [n_j\lambda])/2$, $N_j = (n_j\lambda + [n_j\lambda])/2$, so that $0 \leqq \delta_j < \frac{1}{2}$, and $N_j \geqq [h]$. We have $\sin n_j\lambda\theta - \sin[n_j\lambda]\theta = 2 \sin \delta_j\theta \cos N_j\theta$, and we shall study separately the integral

$$U^* = \int_{\pi/h}^{\pi} \phi(\lambda\theta) \, (\sum_{h}^{2h-1} \alpha_j \sin [n_j\lambda]\theta) \, d\theta/\theta$$

and the difference $U - U^*$.

We have

$$|U^*| \leqq \{ \int_{\pi/h}^{\pi} |\phi(\lambda\theta)/\theta|^q \, d\theta\}^{1/q} \{ \int_{0}^{\pi} |\sum_{h}^{2h-1} \alpha_j \sin [n_j\lambda]\theta|^r \, d\theta\}^{1/r}.$$

Denoting by I, J the two terms of this product, we have

$$I^q = \int_{\pi\lambda/h}^{\pi\lambda} |\phi(t)/t|^q \, dt \cdot \lambda^{q-1} < \lambda^{q-1} \int_{\pi\lambda/h}^{\pi} |\phi(t)/t|^q \, dt.$$

The last integral gives, by a classical process of integration by parts, and owing to $\int_0^{\eta} |\phi(t)|^q \, dt = o(\eta)$, a result which is $o(h/\lambda)^{q-1}$. Hence $I^q = o(h^{q-1})$ and $I = o(h^{1-1/q})$. As for J, since $r \geqq 2$ and all $[n_j\lambda]$ are different, it is less, by the theorem of Hausdorff-Young, than the product of $(\sum |\alpha_j|^q)^{1/q}$

by an absolute constant. Hence

$$| U^* | = o(h^{1-1/q}) \{ \sum_{h}^{2h-1} | \alpha_j |^q \}^{1/q}.$$

Now we consider

$$U - U^* = 2 \sum_{j=h}^{2h-1} \alpha_j \int_{\pi/h}^{\pi} \phi(\lambda\theta) \sin \delta_j \theta \cos N_j \theta \, d\theta/\theta$$

and we shall show that $| U - U^* | = \{ \sum_{h}^{2h-1} | \alpha_j | \} \cdot o(1)$ as $h \to \infty$.

For this purpose we consider the integral

$$W = \int_a^b \phi(\lambda\theta) \sin \delta\theta \cos m\theta \, d\theta/\theta,$$

where $0 < a < b < 2\pi$, $0 < \lambda < 1$, $0 < \delta < \frac{1}{2}$, $m \geqq [h]$ and we proceed to find an uuper bound for $| W |$.

Write

$$W = \int_{a\lambda}^{b\lambda} \phi(t) \sin (\delta/\lambda) t \cos (m/\lambda) t \, dt/t = \int_{a'}^{b'} \phi(t) \sin (\delta/\lambda) t \cos Mt \, dt/t,$$

where $a\lambda = a'$, $b\lambda = b'$, $0 < a' < b' < 2\pi$, and $M = m/\lambda > m$.

Let $\psi(t) = \phi(t)(\sin \delta/\lambda t)/t = \phi(t) g(t)$. Then $\psi(t) \varepsilon L$. As it is well known,

$$| \int_{a'}^{b'} \psi(t) \cos Mt \, dt | \leqq \tfrac{1}{2} \int_{a'}^{b'} | \psi(t + \pi/M) - \psi(t) | \, dt$$
$$+ \tfrac{1}{2} \int_{a'}^{a'+\pi/M} | \psi | \, dt + \tfrac{1}{2} \int_{b'}^{b'+\pi/M} | \psi | \, dt \leqq \tfrac{1}{2} \int_{a'}^{b'} | \psi(t + \pi/M) - \psi(t) | \, dt$$
$$+ \tfrac{1}{2}(\delta/\lambda) \int_{a'}^{a'+\pi/M} | \phi | \, dt + \tfrac{1}{2}(\delta/\lambda) \int_{b'}^{b'+\pi/M} | \phi | \, dt.$$

We have $\int_E | \phi | dt = \xi(| E |)$ where $\xi \to 0$ with $| E |$, measure of E. Thus

$$| \int_{a'}^{b'} \psi(t) \cos Mt \, dt | \leqq \tfrac{1}{2} \int_{a'}^{b'} | \psi(t + \pi/M) - \psi(t) | \, dt + (\delta/\lambda) \xi(\pi/M).$$

Now

$$\psi(t + \pi/M) - \psi(t)$$
$$= [\phi(t + \pi/M) - \phi(t)] g(t + \pi/M) + \phi(t) [g(t + \pi/M) - g(t) |.$$

Now $| g |$ is bounded by δ/λ, and $| dg/dt |$ by $c(\delta/\lambda)^2$ where

$$c = \max | (d/dt)(\sin t/t) |.$$

Hence

$$\tfrac{1}{2}\int_{a'}^{b'}|\psi(t+\pi/M)-\psi(t)|\,dt$$
$$\leqq(\delta/\lambda)\tfrac{1}{2}\int_{a'}^{b'}|\phi(t+\pi/M)-\phi(t)|\,dt+(\pi/2M)c(\delta/\lambda)^2\int_{a'}^{b'}|\phi|\,dt.$$

Denoting by $\omega(\eta)$ the integral modulus of continuity of the given function f we have, collecting results

$$\left|\int_{a'}^{b'}\psi(t)\cos Mt\,dt\right|\leqq(\delta/\lambda)\xi(\pi/M)+(\delta/\lambda)\omega(\pi/M)+C\delta^2/\lambda^2M,$$

C being a constant (depending on the value of $f(x)$ only). Since $\delta<\tfrac{1}{2}$, and $\lambda M=m$, and $M>m\geqq[h]$, we have

$$|W|\leqq C(\xi(\pi/h)+\omega(\pi/h)+1/h)/\lambda=\mu(h)/\lambda,$$

C being independent of a,b,λ,δ,m; and $\mu(h)$ tending to zero as $h\to\infty$, and depending on h only.

Now, if $\lambda>\mu(h)^{\frac{1}{2}}$ then $|W|\leqq C\mu(h)^{\frac{1}{2}}$. On the other hand, if $\lambda<\mu(h)^{\frac{1}{2}}$

$$|W|<\delta\int_0^{2\pi}|\phi(\lambda\theta)|\,d\theta<(1/2\lambda)\int_0^{2\pi\lambda}|\phi(t)|\,dt=o(1)$$

for $h\to\infty$. Hence $|W|=o(1)$ uniformly as $h\to\infty$ and thus, as stated

$$|U-U^*|=\{\sum_h^{2h-1}|\alpha_j|\}\cdot o(1).$$

Since $\sum|\alpha_j|/h\leqq\{\sum|\alpha_j|^q/h\}^{1/q}$, $|U-U^*|=o(h^{1-1/q})\{\sum|\alpha_j|^q\}^{1/q}$, hence, taking into account the result about U^*,

$$|U|=o(h^{1-1/q})\{\sum_h^{2h-1}|\alpha_j|^q\}^{1/q},$$

and this result is obviously valid also for $|U_1|+\cdots+|U_P|$.

To find a final result for S it remains to appraise V, which is done in a classical way:

$$|V|<\int_{\pi/h}^{\pi}|\phi(t)/t|\cdot|\sum_h^{2h-1}\alpha_j\sin n_jt|\,dt$$
$$\leqq\{\int_{\pi/h}^{\pi}|\phi(t)/t|^q\,dt\}^{1/q}\{\int_0^{2\pi}|\sum\alpha_j\sin n_jt|^r\,dt\}^{1/r}$$

$=o(h^{1-1/q})\{\sum_h^{2h-1}|\alpha_j|^q\}^{1/q}$, by the Hausdorff-Young theorem.

Collecting results, we have $S=o(h^{1-1/q})\{\sum_h^{2h-1}|\alpha_j|^q\}^{1/q}$. But it is well known that $T^{1/r}=S$ for some α_j satisfying $\{\sum_h^{2h-1}|\alpha_j|^q\}^{1/q}=1$. Hence

$T^{1/r} = o(h^{1-1/q}) = o(h^{1/r})$, that is to say, $T = o(h)$, as stated. Hence also

$$\sum_{h}^{2h-1} | s_{n_j} - f |^r = o(h).$$

To prove now our theorem it is obvious that it suffices to consider the case $k = 2^{2m} - 1$. Write

$$\sum_{1}^{k} | s_{n_j} - f |^r = \sum_{1}^{2^m - 1} + \sum_{2^m}^{2^{m+1}-1} + \cdots + \sum_{2^{2m-1}}^{2^{2m}-1}.$$

Consider separately the *first* sum; since $2^m - 1 \sim k^{\frac{1}{2}}$, and since $\int_0^\eta | \phi_x(t) | \, dt = o(\eta)$,

$$\sum_{1}^{k^{\frac{1}{2}}} | s_{n_j} - f |^r = o(k^{\frac{1}{2}} \log^r n_k) = o(k^{\frac{1}{2}} \log^r k) = o(k).$$

Now the sum of the other terms is, by the result just proved,

$$o(2^m) + o(2^{m+1}) + \cdots + o(2^{2m-1}) = o(2^{2m}) = o(k).$$

This completes the proof.

It remains now to get rid of our hyppothesis $\epsilon > 2$. This can be done easily by breaking the sequence $\{n_k\}$ in a finite number of complementary sequences, since there exists an integer C such that $(n_{k+C} - n_k)/n_k > 2/k$.

Remarks on the hypothesis about the sequence $\{n_k\}$. It is certain that if n_k increases *too* rapidly with k, the theorem cannot be true any more. (Consider classical examples of a continuous function whose Fourier Series diverges at a point). But it is not excluded that $n_k = O(k^A)$ can be replaced by a better condition. We wish only to prove here that the theorem is not true for $n_k = O(2^k)$, and not even for $n_k = O(2^{k^\alpha})$ if $\alpha > \frac{1}{2}$. More precisely, the theorem is false if n_k increases sufficiently rapidly to have $\log n_k / k^{\frac{1}{2}} \to \infty$.

The proof is rather simple. Consider a continuous function $f(x)$, with period 2π, and its partial sums s_n at the point $x = 0$. One has

$$s_n = s_n(0) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin nt \, dt/t + o(1).$$

Now

$$\begin{aligned} R_k &= (| s_{n_1} | + | s_{n_2} | + \cdots + | s_{n_k} |)/k \\ &\geqq (\phi_1(u) s_{n_1} + \phi_2(u) s_{n_2} + \cdots + \phi_k(u) s_{n_k})/k, \end{aligned}$$

where $\phi_1, \phi_2, \cdots$ are the Rademacher functions. Hence, if we want to prove the unboundedness of R_k it is sufficient to prove the unboundedness—for some $f(t)$—of

$$\int_{-\pi}^{\pi} f(t)\{\phi_1(u)\sin n_1 t + \cdots + \phi_k(u)\sin n_k t\}\,dt/kt = \int_{-\pi}^{\pi} f(t) H_{k,u}(t)\,dt.$$

The kernel $H_{k,u}(t)$ depends on two parameters k, u. It will be enough to show that $\int_{-\pi}^{\pi} |H_{k,u}(t)|\,dt$ cannot be bounded independently of k, u. But, if this integral were bounded, we would have

$$\int_0^{\pi} dt \int_0^1 |H_{k,u}(t)|\,du = O(1)$$

independently of k. We know that

$$\begin{aligned}\int_0^1 |H_{k,u}(t)|\,du &> A(1/kt)(\sin^2 n_1 t + \text{—} \sin^2 n_k t)^{\frac{1}{2}} \\ &> (A/k^{\frac{3}{2}}t)\{|\sin n_1 t| + \cdots + |\sin n_k t|\}.\end{aligned}$$

But the integral of this expression with respect to t, in $0, \pi$, is of order $(1/k^{\frac{3}{2}})(\log n_1 + \cdots + \log n_k)$, hence unbounded if $\log n_k / k^{\frac{1}{2}} \to \infty$.

If we come now the condition of regularity (2) of our theorem, we shall show that *some* condition of this kind is necessary to prove the theorem.

We shall base the proof of this remark on the following lemma

LEMMA. *Consider the integral*

$$J_{n,p} = \int_0^{\pi/p} |\sin(n+1)\theta + \cdots + \sin(n+p)\theta|\,d\theta/p\theta$$

(*in which we think of* n, p, *and* n/p *as being large*)*; this integral is larger than* $A\log(n/p)$, A *being a positive constant.*[2]

Proof. We have

$$\begin{aligned}J_{n,p} &= \int_0^{\pi/p} |\cos(n+p+\tfrac{1}{2})\theta - \cos(n+\tfrac{1}{2})\theta|\,d\theta/\{2p\theta\sin(\theta/2)\} \\ &> \int_0^{\pi/p} |2\sin p(\theta/2)\sin(n+(p+1)/2)\theta|/p\theta^2 \\ &= \int_0^{\pi} |2\sin(t/2)\sin(n/p+\tfrac{1}{2}+1/2p)t|\,dt/t^2 > A\log(n/p).\end{aligned}$$

Take now the sequence n_k in the following way. $n_k = 2^{p\nu}$ for $k = 2^{\nu}$, p being a fixed integer. For $k = 2^{\nu} + h$ where $h = 1, 2, \cdots, 2^{\nu} - 1$ $n_k = 2^{p\nu} + h2^{\nu(p-2)}$.

Obviously $n_k = O(k^p)$.

---

[2] This has certainly been known to many authors.

Let now $f(x)$ be continuous, and s_n be the partial sum of order n for $x=0$.

Let $k=2^\nu$. We have

$$(|s_1|+|s_2+\cdots+|s_{2^{\nu+1}-1}|)/(2^{\nu+1}-1)$$
$$> (|s_{2^\nu}|+\cdots+|s_{2^{\nu+1}-1}|)/(2^{\nu+1}-1).$$

Hence, to prove unboundedness, it is enough to consider the integral

$$\int_{-\pi}^{\pi} f(t)\{\sin k^p t+\sin(k^p+k^{p-2})t+\sin(k^p+2k^{p-2})t$$
$$+\cdots+\sin(k^p+k^{p-1}-k^{p-2})t\}dt/kt,$$

and to prove the unboundedness of

$$\int_0^{\pi}|\sin k^p t+\sin(k^p+k^{p-2})t+\sin(k^p+2k^{p-2})t$$
$$+\cdots+\sin(k^p+k^{p-1}-k^{p-2})t\,|\,dt/kt,$$

or, setting $k^{p-2}t=\theta$, the unboundedness of

$$\int_0^{k^{p-2}\pi}|\sin k^2\theta+\sin(k^2+1)\theta+\cdots+\sin(k^2+k-1)\theta\,|d\theta/\theta,$$

which is $>A\log k$, by the Lemma.

We could as well have taken for $k=2^\nu+h$ and $h=1,2,\cdots,2^\nu-1$, $n_k=2^{p\nu}+[h2^{\nu(p-1-\epsilon)}]$. In all those examples $(n_{k+1}-n_k)k/n_k$ does not remain above a positive number.

THEOREM II. *Let $f(x)$ be continuous, with period 2π, $s_n(x)$ be the partial sum of order n of its Fourier Series. Let M_n denote the maximum of $|s_n(x)|$ for $0\leqq x\leqq 2\pi$. There exists continuous functions for which $(M_1+M_2+\cdots+M_n)/n$ is unbounded.*

Proof. Write, D_n denoting the Dirichlet kernel,

$$T(m,p,x)=(s_{m+1}(x)+s_{m+2}(x)+\cdots+s_{m+p}(x))/p$$
$$=\frac{1}{\pi}\int_0^{2\pi} f(t)\{D_{m+1}(x-t)+\cdots+D_{m+p}(x-t)\}dt/p,$$

and denote by $L(m,p,\theta)$ the kernel $(1/p)[D_{m+1}(\theta)+\cdots+D_{m+p}(\theta)]$.

Let n be given, and k the integer such that $2k^2\leqq n<2(k+1)^2$. One has

$$(M_1+\cdots+M_n)/n$$
$$>(M_1+\cdots+M_{2k^2})/2(k+1)^2>(M_{k^2+1}+\cdots+M_{2k^2})/2(k+1)^2.$$

It will then be sufficient to prove the unboundedness of

$$R_k = (M_{k^2+1} + \cdots + M_{2k^2})/k^2.$$

Let $x_j = j2\pi/k$ $(j = 0, 1, \cdots, k-1)$. Denote by $\phi_0(u), \phi_1(u), \cdots$ the Rademacher functions. Then, obviously, for any value of u,

$$(1/k)(M_{k^2+1} + \cdots + M_{k^2+k}) \geqq \phi_0(u) T(k^2, k, x_0)$$
$$(1/k)(M_{k^2+k+1} + \cdots + M_{k^2+2k}) \geqq \phi_1(u) T(k^2 + k, k, x_1)$$
$$(1/k)(M_{k^2+k(k-1)+1} + \cdots + M_{2k^2}) \geqq \phi_{k-1}(u) T(k^2 + k(k-1), k, x_{k-1}).$$

Hence

$$\begin{aligned} R_k &\geqq (1/k)\{\phi_0(u) T(k^2, k, x_0) + \phi_1(u) T(k^2 + k, k, x_1) \\ &\qquad + \cdots + \phi_{k-1}(u) T(2k^2 - k, k, x_{k-1})\} \\ &= (1/\pi k) \int_0^{2\pi} f(t) \{\phi_0(u) L_{k^2,k}(t - x_0) + \phi_1(u) L_{k^2+k,k}(t - x_1) \\ &\qquad + \cdots + \phi_{k-1}(u) L_{2k^2-k,k}(t - x_{k-1})\} \\ &= \frac{1}{\pi} \int_0^{2\pi} f(t) H_{k,u}(t)\, dt \text{ say.} \end{aligned}$$

The argument already used above shows that it is sufficient to prove the unboundedness of

$$\int_0^{2\pi} dt \int_0^1 |H_{k,u}(t)|\, du > \int_0^{2\pi} (L^2_{k^2,k}(t - x_0) + \cdots)^{\frac{1}{2}} dt/k.$$

But the last integral is larger than

$$k^{-1} \sum_{j=0}^{k-1} \int_{x_j - \pi/k}^{x_j + \pi/k} |L_{k^2+jk,k}(t - x_j)|\, dt = k^{-1} \sum_{j=0}^{k-1} \int_{-(\pi/k)}^{\pi/k} |L_{k^2+jk,k}(\theta)|\, d\theta,$$

and by the last lemma one has

$$\int_{-(\pi/k)}^{\pi/k} |L_{k^2+jk,k}(\theta)|\, d\theta > C \log(k + j) \geqq C \log k.$$

Hence for suitable choice of the bounded function $f(t)$ (depending on k) $R_k \geqq A \log k$, which, by a classical theorem, proves our result.

More precisely, $\omega(n)$ being any function increasing less rapidly than $\log n$, there is a continuous $f(x)$ such that

$$(1/\omega(n)) \cdot (M_1 + M_2 + \cdots + M_n)/n$$

is unbounded.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

BIBLIOGRAPHY.

[1] J. Mazcinkiewicz, "Sur la sommabilité forte des séries de Fourier," *Journal of the London Mathematical Society*, vol. 14 (1939), pp. 162-168.

[2] Z. Zalcwasser, "Sur la sommabilité des séries de Fourier," *Studia Mathematica*, vol. 6 (1936), pp. 82-88.

[3] A. Zygmund, *Trigonometrical Series*, Warszawa-Lwów, 1935, p. 238.

[4] ———, "On the convergence and summability of power series on the circle of convergence (II)," *Proceedings of the London Mathematical Society*, vol. 47 (1942), pp. 326-350.

ERRATA.

P. Hartman and A. Wintner, "Asymptotic integrations of linear differential equations," this JOURNAL, vol. 77, pp. 45-86.

On page 62, formula (67), and on page 72, formula (119), read $\epsilon(a)$ in place of the factor $(\epsilon(j)+\epsilon(a)-1)$ in the exponent of t.

On page 63, formula (70), read $>$ instead of $<$ (twice); in place of the two lines following (70), read "In this case, $y(t)$ is unique."

On page 64, line 11, read [6] in place of [5].

On page 66, formula (80), read $(b(q)-k)$ in place of the factor $(b(q)-h)$.

On page 68, formula (91), read $(\epsilon(a)-\epsilon(j))$ in place of the factor $(\epsilon(a)-1)$ in the exponent of t.

ON THE FORMS OF THE PREDICATES IN THE THEORY OF CONSTRUCTIVE ORDINALS (SECOND PAPER).*

By S. C. Kleene.

In the first paper of this title, the author presented a technique for use in attempts to reduce inductive definitions to explicit definitions, and claimed by means of it to show that the predicates $a \varepsilon O$ and $a <_O b$ of the system S_3 of notation for ordinal numbers are expressible in the respective forms $(x)(Ey)R(a, x, y)$ and $(x)(Ey)S(a, b, x, y)$ where R and S are primitive recursive predicates. To apply the technique, the inductive definition of a predicate P is first rewritten as an equivalence, and then a solution of the latter for P as an unknown predicate is sought.

A fallacy occurs in the preliminary step:[1] the equivalence does not fully express the inductive definition, but may admit other solutions for P than the predicate defined by the inductive definition. It is thus necessary to supplement the application of the technique by a proof that the particular solution obtained is not "extraneous". It appears now that this supplementary reasoning can be given for the applications of the technique in which the proposed solution has only an existential quantifier, but fails in the final application (to $a \varepsilon Q$ in **14**) where there is also a universal quantifier.

In the present paper we shall show that the result claimed in the first paper becomes correct, if a variable α for a 1-place number-theoretic function is substituted for the number variable x; i. e., $a \varepsilon O$ and $a <_O b$ are expressible in the respective forms $(\alpha)(Ey)R(a, \alpha, y)$ and $(\alpha)(Ey)S(a, b, \alpha, y)$ where R and S are primitive recursive (**29** Theorem II).[2] Conversely, every predicate of the form $(\alpha)(Ey)R(a,\alpha,y)$ where R is general recursive is expressible in the form $\xi(a) \varepsilon O$ where ξ is primitive recursive (**25** Theorem I). It follows by the theory of arithmetical predicates and function quantifiers[3] that $a \varepsilon O$ and $a <_O b$ are not expressible in the dual forms $(E\alpha)(y)R(a,\alpha,y)$ and $(E\alpha)(y)S(a,b,\alpha,y)$ with recursive R and S, respectively; a fortiori,[4] they are not arithmetical, so in particular the result claimed originally is false. Most of the work done in the first paper is utilized in proving the present results. We give an example of a primitive recursive linear ordering of the natural numbers which is a well-ordering with respect to the arithmetical functions but not with respect to all the number-theoretic functions (**26** (J)). Classical methods are used in this example and to some extent elsewhere. The results hold also for the relative version of the theory of constructive ordinals (in **30**).

---

* Received December 4, 1953. The references to [13] and [15] were added June 3, 1954. The ideas for the proof of Theorems I and II were obtained in March 1950 while the author was on research leave supported by the University of Wisconsin Graduate School and held a John Simon Guggenheim Memorial Fellowship. The first paper of this title appeared in this Journal, vol. 66 (1944), pp. 41-58.

[1] This is the error noted in the bibliographical references to the first paper in [10] and [11].

[2] A function or predicate with variables for 1-place number-theoretic functions as well as for natural numbers we call *primitive recursive* (as in [12]), if as a function or predicate of its number variables it is primitive recursive uniformly in its function variables in the sense of [11] p. 234. Like terminology has already been used in the case of general recursiveness in [10] pp. 681-682 and in [12].

Some of the numbering and symbolism is continued from the first paper. Sections **18** and **19** indicate in detail what can be kept and what must be rejected in that paper. Readers not having that paper at hand may begin with **20**. For their convenience, information from that paper will be repeated here as needed; and they can choose from various sources cited in the footnotes for the notations and results of the theory of recursive functions which are employed.

18. The reduction technique. On p. 46 of the first paper, we rewrote the inductive definition 1-4 as an equivalence (18), assuming uncritically that the latter is equivalent as a definition of $P(a)$ to the former.

But suppose that the formal deductive system, for which the $P(a)$ of 1-4 is the arithmetized provability predicate, includes among its symbols $\supset$ and among its rules of inference modus ponens

$$\frac{\mathrm{A},\ \mathrm{A} \supset \mathrm{B}}{\mathrm{B}.}$$

Also suppose the arithmetrization to have been carried out (as is usual) so that $A(a)$, $B(a,b)$ and $C(a,b,c)$ are each true only when a is (the Gödel number of) a formula. In this case, (18) can have as solution for $P(a)$ not only $P(a) \equiv \{a \text{ is provable}\}$, which we intended, but also $P(a) \equiv \{a \text{ is a formula}\}$. For each formula A is inferable from two other formulas, e.g., A and $\mathrm{A} \supset \mathrm{A}$, by modus ponens. Hence, when $P(a) \equiv \{a \text{ is a formula}\}$: For each a which makes the left side of (18) true, the third disjunctive member $(Ex)(Ey)[P(x) \& P(y) \& C(a,x,y)]$ of the right side, and hence the right side itself, will be true. Conversely, for each a which makes the right side true, and hence makes either $A(a)$, or $B(a,x)$ for some x, or $C(a,x,y)$ for some x and y, true, the left side will be true.

---

[3] [12] Theorem 3.

[4] [12] Theorem 5.

What is lost in rewriting the inductive definition 1-4 as the equivalence (18) is part of the effect of the extremal clause 4. The predicate $P(a)$ defined by 1-4 (write it now $P_0(a)$) is that predicate $P(a)$ satisfying (18) for which the set $\hat{a}P(a)$ of the a's which make $P(a)$ true is the least. In other words, $P_0(a)$ is the solution of (18) such that, for any solution $P(a)$ of (18),

$$(18.1) \qquad P_0(a) \rightarrow P(a).$$

This implication can be proved by induction, in the form which corresponds to the inductive definition 1-4 of $P_0(a)$.

In particular, then (18.1) holds when $P(a)$ is our solution

$$(23) \qquad P(a) \equiv (Ex)R(a,x)$$

of (18). But also, by course-of-values induction on x (using (13)), we can show from the definition (22) of $R(a,x)$ that

$$(18.2) \qquad (a)[R(a,x) \rightarrow P_0(a)],$$

whence (via $(x)(a)[R(a,x) \rightarrow P_0(a)]$ and $(a)(x)[R(a,x) \rightarrow P_0(a)]$)

$$(18.3) \qquad (Ex)R(a,x) \rightarrow P_0(a).$$

So, by (18.1) (applied to $(Ex)R(a,x)$ as the $P(a)$) and (18.3), our solution $P(a) \equiv (Ex)R(a,x)$ ((23) with (22)) of (18) is in fact $P_0(a)$.

Thus with the help of the supplementary argument given now, we do establish that $\{a \text{ is provable}\}$ is expressible in the form $(Ex)R(a,x)$ with a primitive recursive R.

Likewise, at each place in the paper where we rewrote an inductive definition as an equivalence, we lost part of the effect of the extremal clause. But in every case where we solved that equivalence by a predicate of the form $(Ex)R(a,x)$ or $(Ex)R(a,b,x)$, we can show by induction on x that our solution is the predicate defined by the inductive definition.

19. Conclusions from the first paper. In this manner the conclusions of the first paper can be upheld, except those of **14** and **15** including Theorem 1. and (41) in **16**.[5] By **25** Corollary to Theorem I, we see now that Theorem 1 and (41) are false. By **25** Theorem I with [12] Theorem 5, likewise Theorem IX of [9] Section 15, our "proof" of which depended on Theorem 1, is false.

---

[5] On p. 43 omit the stars on (4) and (11) (cf. [11] *86 and *95 pp. 162-163). On p. 52 after (V) and (VI) the "(Use (IV).)" is unnecessary. On p. 53 in (30) and

20. The predicates $a \varepsilon O$ and $a <_O b$ (recapitulated).[6] In this paper we are writing $T(z, x, y)$ for $T_1(z, x, y)$, and $\Phi(z, x)$ for $\Phi_1(z, x)$ $(\cong U(\mu y\, T_1(z, x, y)))$.[7]

Let $0_O = 1$, $(n+1)_O = 2^{n_O}$. By saying that y defines y_n recursively as a function of n_O, we mean that $(n)\{\Phi(y, n_O)$ is defined and $= y_n\}$. The predicates $a \varepsilon O$ and $a <_O b$ of the system S_3 of notation for ordinal numbers (of the constructive first and second number classes) were introduced by the following inductive definition:[8]

*O*1. $1 \varepsilon O$. *O*2. If $y \varepsilon O$, then $2^y \varepsilon O$ and $y <_O 2^y$. *O*3. If, for each n, $y_n \varepsilon O$ and $y_n <_O y_{n+1}$, and if y defines y_n recursively as a function of n_O, then $3 \cdot 5^y \varepsilon O$ and, for each n, $y_n <_O 3 \cdot 5^y$. *O*4. If $x \varepsilon O$, $y \varepsilon O$, $z \varepsilon O$, $x <_O y$ and $y <_O z$, then $x <_O z$. *O*5. $a \varepsilon O$ and $a <_O b$ only as required by *O*1-*O*4.

We recall the following lemmas (I)-(VIII) from **12**. Here $C(b)$ is a class of numbers depending on b (defined in **12**), such that (as shown in **13**)

$$a \varepsilon C(b) \equiv (Ex) V(a, b, x) \tag{32}$$

for a certain primitive recursive V. We write $a \leqq_O b$ for $a <_O b \vee a = b$, $a >_O b$ for $b <_O a$, etc.

(I) *If* $a \varepsilon O$, *then* $a \neq 0$. (II) *If* $a <_O b$, *then* $b \neq 1$. (III) $C(1)$ *is empty.* (IV) *If* $a <_O b$, *then* $a \varepsilon O$ *and* $b \varepsilon O$. (V) *If* $a <_O 2^y$, *then* $a \leqq_O y$. (VI) *If* $a <_O 3 \cdot 5^y$, *then, for some* n, $a \leqq_O y_n$, *where* $y_n = \Phi(y, n_O)$. (VII) *If* $b \varepsilon O$, *then* $a \varepsilon C(b) \rightarrow a <_O b$. (VIII) *If* $b \varepsilon O$, *then* $a <_O b \rightarrow a \varepsilon C(b)$.

---

(31) read "$b = 3 \cdot 5^y$" for "$a = 3 \cdot 5^y$". On p. 54 in (35) read "$a = 3 \cdot 5^{(a)_3}$" for "$e = 3 \cdot 5^{(a)_3}$" (an adaptation of **14** is used in **27**). In order that **16** Theorem 2 give the full force of (41)-(44), it should be added in each part that the function also map one-to-one the complement of the one set on a subset of the complement of the other; this being added, the first part becomes false. Theorem 3 (which should have been indicated as classical) is now included in the much stronger Corollary to Theorem I (also classical).

[6] The reader may wish to follow the modifications outlined in **30** for the relative version of the theory as he reads **20**-**29** giving the absolute version.

[7] **10**; or [9] Sections 4, 7, 10; or [11] pp. 225, 279, 281, 288, 327, 330, 340.

[8] The theory of constructive ordinals in various versions appears in Church-Kleene [4], Church [3], Kleene [8] (cf. Section 3 for n_O, Section 5 for S_3) and Turing [20] Section 7.

We add several further lemmas. Some of these are proved, as were preceding ones, by induction in a form which corresponds to the inductive definition $O1$-$O5$. This is illustrated by the proofs of (VII) in **12** and (XX) in **22**, which are given in full.

(IX) *If* $a \varepsilon O$, *then* $a \geqq_O 1$. (By induction.) (X) *If* $a \varepsilon O$, *then, for each number-theoretic function* $\alpha(n)$ *such that* $\alpha(0) = a$ *and* $(n)[\alpha(n) \neq 1 \rightarrow \alpha(n) >_O \alpha(n+1)]$, *there is a number* k *such that* $\alpha(k) = 1$. (By induction. For example in the part where it is assumed that $a = 2^y$ with $y \varepsilon O$, by (I), $a \neq 1$, so $a = \alpha(0) >_O \alpha(1)$. By (V), either $y = \alpha(1)$ (Case A) or $y >_O \alpha(1)$ (Case B). We apply the hypothesis of the induction in Case A to $\beta(n) = \alpha(n+1)$; in Case B (then by (II), $y \neq 1$) to $\beta(n) = y$ if $n = 0$, $= \alpha(n)$ if $n \neq 0$.)

(XI) *For each* a, $\overline{a <_O a}$. (By reductio ad absurdum, from (II), (IV), and (X) applied to $\alpha(n) = a$.) (XII) *If* $a <_O b$, *then* $\overline{b <_O a}$. (By (IV), $O4$ and (XI).) (XIII) If $c \varepsilon O$, $a \leqq_O c$ *and* $b \leqq_O c$, *then* $a <_O b$ *or* $a = b$ *or* $a >_O b$. (By induction on c.)

Together $O4$, (XI) and (XII) give that O is partially ordered by $<_O$. Then by (XIII) the members of O which are $<_O$ a given member b are linearly ordered, and indeed by (X) with (II) well-ordered, by $<_O$. Thus far we have not had to mention the classical second number class.

Letting $|1| = 0$, $|2^y| = |y| + 1$ for $y \varepsilon O$, and $|3 \cdot 5^y| = \lim_n |y_n|$ for $3 \cdot 5^y \varepsilon O$ and $y_n = \Phi(y, n_O)$, each member b of O represents an ordinal $|b|$ of the first or second number class, in such a manner that whenever $a <_O b$ then $|a| < |b|$ and that for each $\alpha < |b|$ there is a member a of O such that $|a| = \alpha$ and $a <_O b$.[9] By this with (XIII), the ordering of the set $\hat{a}(a <_O b)$ by $<_O$ is of the type $|b|$. The least ordinal $> |b|$ for every $b \varepsilon O$ is called ω_1.

21. A digression concerning recursive well-orderings. A partial recursive function $\mathrm{enm}(b, m)$ which for each fixed $b >_O 1$ enumerates the numbers $<_O b$ is constructible by the method used in the earliest version of the theory.[10] However it is worth noting that the result is immediate from **12** and **13**. It is easily arranged that enm be primitive recursive and that the enumeration be without reptitions when b represents an infinite ordinal. In detail:

---

[9] Cf. **11**, [3], [8], and for an earlier version [4]. Also for (XIII), cf. [3] pp. 229-230.

[10] [4] p. 20 enm, and [8] p. 154 end Section 3. Also cf. [20] pp. 188-189 Inf.

Combining (VII) and (VIII) with (32),

$$(21.1) \qquad b \,\varepsilon\, O \rightarrow (a <_O b \equiv (Ex)V(a,b,x)).$$

Let $\mathrm{Fin}(b) \equiv (En)_{n<b}[b = n_O]$, so $\mathrm{Fin}(b) \equiv \{b \,\varepsilon\, O$ by $O1$ and $O2$ only$\}$ $\equiv \{b$ represents a finite ordinal$\}$. Let $\mathrm{nat}(b) = \mu n_{n<b}[b = n_O]$, so that $\mathrm{nat}(n_O) = n$. (When $\mathrm{Fin}(b)$, $\mathrm{nat}(b)$ is $|b|$ as a natural number.) We use $(x)_i$ in the sense of [11].[11] Now in the case $\overline{\mathrm{Fin}}(b)$ (whether or not $b \,\varepsilon\, O$), let $\mathrm{enm}(b,m)$ be $(m)_0$, if $V((m)_0, b, (m)_1)$ and $(m)_0$ is not among $\mathrm{enm}(b,0), \cdots, \mathrm{enm}(b, m-1)$, and otherwise let $\mathrm{enm}(b,m)$ be the first among $0_O, \cdots, m_O$ which is not among $\mathrm{enm}(b,0), \cdots, \mathrm{enm}(b,m-1)$. In the case $\mathrm{Fin}(b)$, let $\mathrm{enm}(b,m) = (\mathrm{nat}(b) \dot{-} (m+1))_O$.[12] This gives a course-of-values recursion on m for $\mathrm{enm}(b,m)$, so enm is primitive recursive.[13]

(XIV) *If* $b \,\varepsilon\, O$ *and* $\overline{\mathrm{Fin}}(b)$, *then* $(n)[n_O <_O b]$. (By induction.)
(XV) *There is a primitive recursive function* $\mathrm{enm}(b,m)$ *such that*: (i) *If* $b = n_O$, *then* $\mathrm{enm}(b,m) = (n \dot{-} (m+1))_O$. (ii) *If* $b \,\varepsilon\, O$ *and* $\overline{\mathrm{Fin}}(b)$, *then* $\mathrm{enm}(b,0)$, $\mathrm{enm}(b,1)$, $\mathrm{enm}(b,2), \cdots$ *is an enumeration without repetitions of* $\hat{a}(a <_O b)$.

In Cantor's theory, ordinals arise on the one hand as the order types of well-ordered sets, and on the other as the objects generated by certain principles. In the present theory of constructive ordinals, it is the latter version of the theory which is made constructive. We inquire now into the relation between this and the result of making the first version constructive, i. e., of requiring the well-ordering to be recursive.

(A) *There is a partial recursive ternary relation* $m <_b n$ *such that, for each fixed* b: *If* $b \,\varepsilon\, O$ *and* $\overline{\mathrm{Fin}}(b)$, *then the binary relation* $m <_b n$ *is a general recursive well-ordering of the natural numbers of order type* $|b|$.[14]

Proof. Suppose $b \,\varepsilon\, O$ and $\overline{\mathrm{Fin}}(b)$. The numbers $<_O b$ are well-ordered by $<_O$ with the order type $|b|$ (end **20**). But $\mathrm{enm}(b,0), \mathrm{enm}(b,1), \mathrm{enm}(b,2), \cdots$ is an enumeration of them without repetitions. So it will suffice to define

---

[11] [11] p. 230, viz., $(x)_i = \{$the number of times x contains the $i+1$-st prime number as factor$\}$ if $x \neq 0$, $= 0$ if $x = 0$. The author originally introduced $(x)_1, \cdots, (x)_m$ in [9] (1943) p. 50 for an unspecified m-tuple of functions with certain properties, and then in the first paper of this title (1944) 5 defined $(x)_i$ so that the $(x)_{i+1}$ there is the $(x)_i$ now; but since 1946 he has undertaken to standardize the notation $(x)_i$ in its present meaning (e. g., it was so used at his suggestion in Nelson's [16] (1947)).

[12] $n \dot{-} m = n - m$ if $n \geqq m$, $= 0$ otherwise (cf. [11] p. 223 #6).

[13] Péter [17] Section 1 or [18] Section 3, Kleene [11] Section 46.

[14] A related result is given by Turing [20] p. 183 (vi).

$m <_b n$ partial recursively so that, for the b in question, $m <_b n \equiv \mathrm{enm}(b,m) <_O \mathrm{enm}(b,n)$. To determine whether $\mathrm{enm}(b,m) <_O \mathrm{enm}(b,n)$, we first check that $m \neq n$. If $m \neq n$, then either $\mathrm{enm}(b,m)$ occurs in the enumeration $\mathrm{enm}(\mathrm{enm}(b,n),0), \mathrm{enm}(\mathrm{enm}(b,n),1), \mathrm{enm}(\mathrm{enm}(b,n),2), \cdots$, which happens if $\mathrm{enm}(b,m) <_O \mathrm{enm}(b,n)$ and only then (since $\mathrm{enm}(b,m) = \mathrm{enm}(b,n) = 1$ is excluded), or else $\mathrm{enm}(b,n)$ occurs in the enumeration $\mathrm{enm}(\mathrm{enm}(b,m),0)$, $\mathrm{enm}(\mathrm{enm}(b,m),1), \mathrm{enm}(\mathrm{enm}(b,m),2), \cdots$, which happens if and only if $\mathrm{enm}(b,m) >_O \mathrm{enm}(b,n)$. So we take

$$m <_b n \cong m \neq n \;\&\; \mathrm{enm}(\mathrm{enm}(b,n),T) = \mathrm{enm}(b,m)$$

where

$$T \cong \mu t[\mathrm{enm}(\mathrm{enm}(b,n),t) = \mathrm{enm}(b,m) \vee \mathrm{enm}(\mathrm{enm}(b,m),t) = \mathrm{enm}(b,n)].$$

(B) *To any general recursive well-ordering of the natural numbers, there is a primitive recursive well-ordering of the natural numbers with the same order type.*[15]

Proof. The construction which follows will lead from a given general recursive well-ordering $\prec$ of order type α to a primitive recursive one $\prec'$ of type $\omega + \alpha$. But for $\alpha \geqq \omega^2$, $\omega + \alpha = \alpha$; and for $\alpha < \omega^2$, a primitive recursive well-ordering of type α is easily defined.

Let e be a Gödel number of (or define recursively) the general recursive ordering predicate $m \prec n$; so $(m)(n)(E!x)T_2(e,m,n,x)$ and

$$T_2(e,m,n,x) \rightarrow [m \prec n \equiv U(x) = 0].^{16}$$

For any natural number n, let the *image of n* be $2^n \cdot 3^{\max_{m \leqq_n} \mu x T_2(e,m,n,x)}$. If l is the image of some number (in symbols, $\mathrm{Im}(l)$), that number is $(l)_0$. In the new well-ordering $\prec'$, the non-images in order of magnitude among themselves shall precede the images ordered among themselves as the numbers of which they are images were ordered in the original ordering $\prec$. Since the class of the non-images is infinite, the new ordering $\prec'$ is of order type $\omega + \alpha$. The following formulas show that the new ordering $\prec'$ is primitive recursive.

$$\begin{aligned}\mathrm{Im}(l) \equiv l = 2^{(l)_0} \cdot 3^{(l)_1} \;\&\; (Em)_{m \leqq (l)_0} T_2(e,m,(l)_0,(l)_1) \\ \&\; (m)_{m \leqq (l)_0} (Ex)_{x \leqq (l)_1} T_2(e,m,(l)_0,x),\end{aligned}$$

---

[15] By [11] pp. 285-287 Example 1, 'primitive recursive' can be strengthened here to 'elementary' in the sense of P. Csillag and L. Kalmár (cf. [6], [18] Section 8, or [11] p. 285).

[16] [9] pp. 52-53 Theorem IV, [11] p. 288 Theorem IX, or (briefly) 10. "$(E!x)$" means "there exists a unique x such that".

$$
\begin{aligned}
k \prec' l \equiv {} & \{\mathrm{Im}(k) \,\&\, \overline{\mathrm{Im}}(l) \,\&\, k < l\} \vee \{\overline{\mathrm{Im}}(k) \,\&\, \mathrm{Im}(l)\} \\
& \vee \{\mathrm{Im}(k) \,\&\, \mathrm{Im}(l) \,\&\, \{[(k)_0 < (l)_0 \,\&\, U(\mu x_{x<l} T_2(e, (k)_0, (l)_0, x)) = 0] \\
& \vee [(k)_0 > (l)_0 \,\&\, U(\mu x_{x<k} T_2(e, (l)_0, (k)_0, x)) \neq 0]\}\}.
\end{aligned}
$$

(C) *There is a primitive recursive ternary relation* $m <_b' n$ *such that, for each fixed* b*: If* $b \varepsilon O$ *and* $\overline{\mathrm{Fin}}(b)$*, then the binary relation* $m <_b' n$ *is a primitive recursive well-ordering of the natural numbers of order type* $|b|$.

Proof. We apply the method of proof of (B) to the $m <_b n$ of (A) as the $m \prec n$. But now we can distinguish primitive recursively which elements in the given ordering are infinite; namely $\{n \text{ is infinite}\} \equiv \overline{\mathrm{Fin}}(\mathrm{enm}(b, n))$. So by letting all numbers not images of infinite elements precede the images only of infinite elements, we avoid the increase for $\alpha < \omega^2$ from α to $\omega + \alpha$. As a Gödel number e of the ordering $m <_b n$, we have $S_2^1(d, b)$,[17] where d is a Gödel number of $m <_b n$ as predicate of b, m, n.

By (A), the order types of general (and hence by (B), or directly by (C), of primitive) recursive well-orderings of the natural numbers include all transfinite ordinals $< \omega_1$. That they include only such ordinals was shown by Markwald [15] (and independently by Clifford Spector, from whom the author first learned the result in February 1954 shortly before [15] was received in Madison).

22. Sums of ordinal notations. The addition of ordinal notations was treated briefly by Church and Kleene.[18] We redo the treatment, in order to secure an additional property of the sum function which we shall need now ((XX) below).

We seek a recursive function $a +_O b$ such that

$$
(22.1) \qquad a +_O b \simeq \begin{cases} a \text{ if } b = 1 \,\&\, a \neq 0 & \text{(Case 1)}, \\ 2^{a +_O y} \text{ if } b = 2^y \text{ where } y \neq 0 & \text{(Case 2)}, \\ 3 \cdot 5^{d_{a,y}} \text{ if } b = 3 \cdot 5^y & \text{(Case 3)}, \\ 7 \text{ otherwise} & \text{(Case 4)}, \end{cases}
$$

where (in Case 3) $d_{a,y}$ defines recursively $a +_O \Phi(y, n_O)$ as a function of n_O.[19] We obtain such a function $a +_O b$ as follows. Let

$$\theta(z, a, y, n) \simeq \Phi_2(z, a, \Phi(y, n)).^7$$

---

[17] End **10**, or [8] p. 153, or [11] p. 342 Theorem XXIII.[23]

[18] [4] p. 18, [8] p. 154. Also Turing [20] p. 188.

[19] Here, since $a +_O \Phi(y, n_O)$ may be undefined, we mean that $(n)\{\Phi(d_{a,y}, n_O) \simeq a +_O \Phi(y, n_O)\}$.

Let p define $\theta(z, a, y, n)$ recursively.[16] Let[11,17]

$$
(22.2) \qquad \begin{aligned}
\psi(z, a, b) \cong \mu t[&\{b = 1 \,\&\, a \neq 0 \,\&\, t = a\} \\
&\vee \{b = 2^{(b)_0} \,\&\, (b)_0 \neq 0 \,\&\, t = 2^{\Phi_2(z, a, (b)_0)}\} \\
&\vee \{b = 3 \cdot 5^{(b)_2} \,\&\, t = 3 \cdot 5^{S_1^3(p, z, a, (b)_2)}\} \\
&\vee \{(b \neq 1 \vee a = 0) \,\&\, (b \neq 2^{(b)_0} \vee (b)_0 = 0) \,\&\, b \neq 3 \cdot 5^{(b)_2} \,\&\, t = 7\}].
\end{aligned}
$$

By the recursion theorem,[20] there is a number e which defines $\psi(e, a, b)$ recursively. Let $a +_O b \cong \psi(e, a, b)$. This gives $a +_O b$ as a partial recursive function, but in fact it is primitive recursive.[21]

(XVI) *If* $a \, \varepsilon \, O$ *and* $b \, \varepsilon \, O$, *then* (a) $a +_O b \, \varepsilon \, O$ *and* (b)

$$(c)[c <_O b \rightarrow a +_O c <_O a +_O b].$$

The proof is by induction on b with a fixed. Parts 1 and 2 (for $b \, \varepsilon \, O$ by $O1$ and $O2$, respectively) are straightforward. For 3, assume that $b = 3 \cdot 5^y$ where y is as in $O3$, and as hypothesis of the induction that (XVI) applies to each y_n as the b. (a) By Case 3 of (22.1), $a +_O b = 3 \cdot 5^d$ where for each n, $\Phi(d, n_O)$ is defined (call it d_n) and $d_n = a +_O y_n$. Applying the hypothesis of the induction, we have by (a) (for y_n as the b) $d_n \, \varepsilon \, O$ and by (b) (for y_n, y_{n+1} as the c, b) $d_n <_O d_{n+1}$. Hence by $O3$, $3 \cdot 5^d \, \varepsilon \, O$ and $d_n <_O 3 \cdot 5^d$, i.e., $a +_O b \, \varepsilon \, O$ and $a +_O y_n <_O a +_O b$. (b) Assume $c <_O b$. By (VI), for some n, either $c = y_n$ (Case A) or $c <_O y_n$ (Case B). Case A. $a +_O c = a +_O y_n <_O a +_O b$. Case B. By the hypothesis of the induction (applied to c, y_n as the c, b), $a +_O c <_O a +_O y_n$. But $a +_O y_n <_O a +_O b$. By $O4$, $a +_O c <_O a +_O b$.

(XVII) *If* $a \, \varepsilon \, O$ *and* $b >_O 1$, *then* $a +_O b >_O a$. (Use (IV), (XVI) (b), (I), (22.1).) (XVIII) *If* $a \, \varepsilon \, O$, *then* $a +_O 1_O >_O 1$. (By (22.1), (I), $O2$ and (IX).) (XIX) *If* $a \, \varepsilon \, O$ *and* $b \, \varepsilon \, O$, *then* $a +_O b \geqq_O a$. (Use (IX), (XVII), (22.1) and (I).)

(XX) *If* $a +_O b \, \varepsilon \, O$, *then* $a \, \varepsilon \, O$ *and* $b \, \varepsilon \, O$. *In fact, for every* c *which* $\varepsilon \, O$: (a) *If* $a +_O b = c$, *then* $a \, \varepsilon \, O$ *and* $b \, \varepsilon \, O$. (b) *If* $a +_O b_1 = a +_O b_2 = c$, *then* $b_1 = b_2$. (c) *If* $b \neq 1 \,\&\, a \leqq_O f <_O a +_O b = c$, *then*

$$(Eg)[g <_O b \,\&\, f = a +_O g].$$

Proof is by induction on c, as follows.

---

[20] [S] p. 153 the last result of Section 2, or [11] p. 352 Theorem XXVII.

[21] For upon using the resulting value $S_1^3(p, e, a, y)$ of $d_{a,y}$ in (22.1), we have a course-of-values recursion for $a +_O b$.[13]

1. Assume that $c = 1$. (a) Assume $a +_O b = c$. Now Cases 3 and 4 of (22.1) are excluded. Also Case 2 is excluded, since this could only apply if $a +_O y = 0$, and in none of the cases of (22.1) (for y as the b) can $a +_O y = 0$. So Case 1 applies, with $a = b = 1$. But $1 \varepsilon O$. So $a, b \varepsilon O$. (b) Assume $a +_O b_1 = a +_O b_2 = c$. By (a) applied to $a + b_1$ and $a + b_2$, $a = b_1 = b_2 = 1$. So $b_1 = b_2$. (c) By (II), $f <_O a +_O b = c$ cannot hold when $c = 1$.

2. Assume that $c = 2^d$ where $d \varepsilon O$, and as hypothesis of the induction that (XX) applies to d as the c. By $O2$, $d <_O c$. (a) Assume $a +_O b = c$. Now only Case 1 or Case 2 of (22.1) can apply. In Case 1, $a = c$ and $b = 1$. But $c, 1 \varepsilon O$. So $a, b \varepsilon O$. In Case 2, $d = a +_O y$ and $b = 2^y$. By (a) of the hypothesis of the induction, then $a, y \varepsilon O$, so using $O2$, $a, b \varepsilon O$ (and $y <_O b$). (b) Assume $a +_O b_1 = a +_O b_2 = c$. If Case 1 applies to both $a +_O b_1$ and $a +_O b_2$, we have $b_1 = b_2 = 1$. If Case 2 applied to $a +_O b_1$ but Case 1 to $a +_O b_2$, we would have $b_1 = 2^y$ where $y \neq 0$ so $b_1 \neq 1$ (and by (a) $a, b_1 \varepsilon O$, and by (IX) $b_1 >_O 1$), but $a = c$. Thus we would have $a = c = a +_O b_1 >_O a$ (by (XVII)), contradicting (XI). So this pair of cases cannot apply; and similarly vice versa. If Case 2 applies to both $a +_O b_1$ and $a +_O b_2$, we have $b_1 = 2^{y_1}$, $b_2 = 2^{y_2}$ and $c = 2^d = 2^{a +_O y_1} = 2^{a +_O y_2}$, whence $d = a +_O y_1 = a +_O y_2$. By the hypothesis of the induction, $y_1 = y_2$. So $b_1 = b_2$. (c) Assume $b \neq 1 \mathbin{\&} a \leqq_O f <_O a +_O b = c$. Since $b \neq 1$, Case 1 does not apply to $a +_O b$; and by Case 2, $d = a +_O y$ and $b = 2^y$, whence using (a) $y \varepsilon O$ and $y <_O b$. Now by (V), either $f = d$ (Case A) or $f <_O d$ (Case B). In Case A, y is a g. In Case B, $y \neq 1$, as otherwise we would have $f <_O a +_O 1 = a$ (since by (a) $a \varepsilon O$, so by (I) $a \neq 0$), contradicting the hypothesis $a \leqq_O f$ (by (XI), (XII)). By the hypothesis of the induction there is an h such that $h <_O y \mathbin{\&} f = a +_O h$. By $O4$ (since $y <_O b$), $h <_O b$. Thus h is a g.

3. Assume that $c = 3 \cdot 5^d$, where, for each n, $\Phi(d, n_O)$ is defined (call it d_n), $d_n \varepsilon O$ and $d_n <_O d_{n+1}$, and as hypothesis of the induction that (XX) applies to each d_n as the c. By $O3$, $d_n <_O c$. (a) Assume $a +_O b = c$. Only Case 1 or 3 of (22.1) can apply. In Case 1, $a = c$ and $b = 1$, so $a, b \varepsilon O$. In Case 3, $b = 3 \cdot 5^y$ and, for each n, $\Phi(d, n_O) \simeq a +_O \Phi(y, n_O)$. But since $\Phi(d, n_O)$ is defined ($= d_n$), $a +_O \Phi(y, n_O)$ and hence $\Phi(y, n_O)$ is defined (call it y_n),[22] so that $d_n = a +_O y_n$. By (a) of the hypothesis of the induction, $a, y_n \varepsilon O$. By (XIX), $a \leqq_O d_n$. Were $y_{n+1} = 1$, we would have $d_{n+1} = a \leqq_O d_n$,

---

[22] Under the weak sense ([8] p. 152 ll. 20-22, or [11] p. 327), $a +_O \Phi(y, n_O)$ can be defined only when $\Phi(y, n_O)$ is defined.

contradicting $d_n <_O d_{n+1}$. Thus $y_{n+1} \neq 1$. Also $d_n <_O d_{n+1} = a +_O y_{n+1}$. Using the hypothesis of the induction to apply (c) (with d_n, d_{n+1}, y_{n+1} as the f, c, b), there is a g such that $g <_O y_{n+1} \,\&\, d_n = a +_O g$. But $d_n = a +_O y_n$. Using the hypothesis of the induction to apply (b) (with d_n as the c), $g = y_n$. Hence $y_n <_O y_{n+1}$. Now by $O3$ we can conclude that $3 \cdot 5^y \, \varepsilon \, O$ (and for each n, $y_n <_O 3 \cdot 5^y = b$). Thus $a, b \, \varepsilon \, O$. (b) Assume $a +_O b_1 = a +_O b_2 = c$. If Case 1 applies to both $a +_O b_1$ and $a +_O b_2$, or Case 3 to either and Case 1 to the other, we can argue as under Part 2. If Case 3 applies to both, then $b_1 = 3 \cdot 5^{y_1}$ and $b_2 = 3 \cdot 5^{y_2}$ for some numbers y_1 $(= (b_1)_2)$ and y_2 $(= (b_2)_2)$, and by the actual construction of the function $+_O$, $a +_O b_1 = 3 \cdot 5^{S_1^3(p,c,a,(b_1)_2)}$ and $a +_O b_2 = 3 \cdot 5^{S_1^3(p,c,a,(b_2)_2)}$. Hence $S_1^3(p, e, a, (b_1)_2) = S_1^3(p, e, a, (b_2)_2)$. By the construction of the function $S_n^m(z, y_1, \cdots, y_m)$,[23] now $(b_1)_2 = (b_2)_2$, i.e., $y_1 = y_2$, so $b_1 = b_2$. (c) Assume $b \neq 1 \,\&\, a \leqq_O f <_O a +_O b = c$. Case 1 does not apply; and by Case 3, $b = 3 \cdot 5^y$ etc. (cf. under (a)). By (VI), for some n, either $f = d_n$ (Case A) or $f <_O d_n$ (Case B). The proof that there is a g is completed in the same manner as in Part 2.

23. Finite sums. For any number-theoretic function $\psi(x)$, let

$$(23.1) \qquad \begin{cases} \sum\limits_{x<0}{}_O \, \psi(x) = 1, \\ \sum\limits_{x<m+1}{}_O \, \psi(x) = \sum\limits_{x<m}{}_O \, \psi(x) +_O \psi(m). \end{cases}$$

Then $\sum\limits_{x<m}{}_O \, \psi(x)$ is a function of m primitive recursive in ψ. Similarly with $\psi(x, a_1, \cdots, a_n)$ instead of $\psi(x)$.

(XXI) *If* $(x)_{x<m}[\psi(x) \, \varepsilon \, O]$, *then* $\sum\limits_{x<m}{}_O \, \psi(x) \, \varepsilon \, O$. (By induction on m, using (XVI) (a).) Conversely: (XXII) *If* $\sum\limits_{x<m}{}_O \, \psi(x) \, \varepsilon \, O$, *then* $(x)_{x<m}[\psi(x) \, \varepsilon \, O]$. (Similarly, using (XX).) (XXIII) *If* $(x)_{x<m}[\psi(x) \, \varepsilon \, O]$ *and* $\psi(m) >_O 1$, *then* $\sum\limits_{x<m}{}_O \, \psi(x) <_O \sum\limits_{x<m+1}{}_O \, \psi(x)$. (Use (XVII) and (XXI).)

24. Securability. We can *represent* a finite sequence $\alpha(0), \cdots, \alpha(x-1)$ of x numbers $(x \geqq 0)$ by the single number $\bar{\alpha}(x) = \prod\limits_{i<x} p_i^{\alpha(i)+1}$.[24] There is a particular primitive recursive predicate $T_1^1(v, z, a)$ such that, for each general recursive predicate $R(a, \alpha, x)$,[2] there is a number f such that

---

[23] [11] p. 342. Erratum (1952 printing): l. 7 from below, replace "$d_* \cdot \cdot \cdot$" by "$d_*[2 \exp \cdot \cdot \cdot]$".

[24] As in [10] p. 680 and [12] Footnote 2; p_i = {the $i + 1$-st prime number} ([11] p. 230 #18).

$(Ex)R(a, \alpha, x) \equiv (Ex)T_1^1(\bar{\alpha}(x), f, a)$.[25] Thus any predicate of the form $(Ex)R(a, \alpha, x)$ with R general recursive can be taken to be of the form $(Ex)R(a, \bar{\alpha}(x))$ with R primitive recursive.

Define $\mathrm{Seq}(w) \equiv w \neq 0 \,\&\, (i)_{i<\mathrm{lh}(w)}[(w)_i \neq 0]$,[25] so $\mathrm{Seq}(w)$ is primitive recursive and $\mathrm{Seq}(w) \equiv \{w = \bar{\gamma}(x)$ for some γ and $x\} \equiv \{w$ represents a finite sequence $\gamma(0), \cdots, \gamma(x-1)\}$. Here $\gamma(i) = (w)_i \dot{-} 1$ $(i = 0, \cdots, x-1)$.

Consider any predicate $R(a, v)$ and any such number $w = \bar{\gamma}(x)$.

We say w (or the x-tuple $\gamma(0), \cdots, \gamma(x-1)$ which w represents) is *secured* (with respect to R and a), if $(Et)_{t \leq x}R(a, \bar{\gamma}(t))$.

We say w (or $\gamma(0), \cdots, \gamma(x-1)$) is *immediately secured* (with respect to R and a), if $\gamma(0), \cdots, \gamma(x-1)$ but no proper initial segment of it is secured, i.e., if $R(a, \bar{\gamma}(x)) \,\&\, (t)_{t<x}\bar{R}(a, \bar{\gamma}(t))$.

We say w (or $\gamma(0), \cdots, \gamma(x-1)$) is *past secured* (with respect to R and a), if w is secured but not immediately so, i.e., if $(Et)_{t<x}R(a, \bar{\gamma}(t))$.

We say w (or $\gamma(0), \cdots, \gamma(x-1)$) is *securable* (with respect to R and a), if w is secured or $(\alpha)(Et)R(a, w * \bar{\alpha}(t))$, i.e., if no matter how $\gamma(0), \cdots, \gamma(x-1)$ (if it is not already secured) is extended as $\gamma(0), \cdots, \gamma(x-1), \alpha(0), \alpha(1), \cdots$, a sequence $\gamma(0), \cdots, \gamma(x-1), \alpha(0), \cdots, \alpha(t-1)$ will eventually result which is secured.[26]

Applying this notion to the case $w = 1 = \bar{\gamma}(0)$, $(\alpha)(Ex)R(a, \bar{\alpha}(x)) \equiv \{1$ is securable with respect to R and $a\}$.

Any number w not of the form $\bar{\gamma}(x)$, i.e., such that $\mathrm{Seq}(w)$, we take to be not secured, immediately secured, past secured or securable.

Now let "securable$_{\mathrm{E}}$" express the notion of securability as just defined explicitly, and "securable$_{\mathrm{I}}$" the notion we define next by an inductive definition *S*1-*S*3. For brevity we are omitting the words "with respect to R and a".

*S*1. If w is secured, then w is *securable*. *S*2. If $\mathrm{Seq}(w)$, and w is not secured,[27] but, for every natural number s, $w * 2^{s+1}$ is *securable*, then w is *securable*. *S*3. w is *securable* only as required by *S*1-*S*2.

(D) *If w is securable$_{\mathrm{I}}$, then w is securable$_{\mathrm{E}}$.*

---

[25] From [11] Theorem IV* pp. 281, 292 (with pp. 231, 291, where $\bar{\alpha}(x)$ and $T_1^1(v, z, a, x)$ are defined) by putting

$$T_1^1(v, z, a) \equiv T_1^1(\Pi_{i<\mathrm{lh}(v)} p_i \exp (v)_i \dot{-} 1, z, a, \mathrm{lh}(v)),$$

noting that $\mathrm{lh}(\bar{\alpha}(x)) = x$ p. 230 #20 (as in [10], or [12] Footnote 2).[11,12,24]

[26] $w * \bar{\alpha}(t)$ represents the latter sequence; cf. [11] p. 230 #21.

[27] We could omit "w is not secured" in *S*2, but we prefer to keep it, so that the inductive definition analyzes securability into two exclusive cases *S*1 and *S*2.

This is easily proved by induction in the form corresponding to the inductive definition $S1$-$S3$. Conversely:

(E) *If w is securable$_{\mathrm{E}}$, then w is securable$_{\mathrm{I}}$.*

We establish this classically. Brouwer argues (in substance) that it is the case intuitionistically (in [2] Footnote 7).

Suppose that $\mathrm{Seq}(w)$ but w is not securable$_{\mathrm{I}}$. Then by $S1$, w is not secured (so $\bar{R}(a, w * \bar{\alpha}(0)))$, and by $S2$ for some number s_0, $w * 2^{s_0+1}$ is not securable$_{\mathrm{I}}$ (put $\alpha(0) = s_0$, so $2^{s_0+1} = \bar{\alpha}(1)$). Then by $S1$, $w * \bar{\alpha}(1)$ is not secured (so $\bar{R}(a, w * \bar{\alpha}(1)))$, and by $S2$ for some number s_1, $w * \bar{\alpha}(1) * 2^{s_1+1}$ is not securable$_{\mathrm{I}}$ (put $\alpha(1) = s_1$, so $\bar{\alpha}(1) * 2^{s_1+1} = \bar{\alpha}(2)$). Then by $S1$, $w * \bar{\alpha}(2)$ is not secured (so $\bar{R}(a, w * \bar{\alpha}(2)))$, etc. Thus we obtain a function α such that $(t)\bar{R}(a, w * \bar{\alpha}(t))$. Hence $\overline{(\alpha)}(Et)R(a, w * \bar{\alpha}(t))$, besides which w is not secured, i.e., w is not securable$_{\mathrm{E}}$.

The remarks in this section apply similarly with $a_1, \cdots, a_n$ for any fixed $n \geqq 0$ in place of a (then $T_1{}^1$ becomes $T_n{}^1$).

25. Classification of $a \varepsilon O$ and $a <_O b$ from below.

THEOREM I. *Given any general recursive predicate $R(a, \alpha, x)$, a primitive recursive function $\xi(a)$ can be found such that $(\alpha)(Ex)R(a, \alpha, x) \equiv \xi(a) \varepsilon O$.*[28]

Proof. By **24**, after choosing a primitive recursive $R(a, v)$ so that $(Ex)R(a, \alpha, x) \equiv (Ex)R(a, \bar{\alpha}(x))$, it will suffice to choose ξ so that $\xi(a) \varepsilon O$ if and only if 1 is securable with respect to this R and a. We begin by defining $\xi(a, w)$ so that $\xi(a, w) \varepsilon O$ if and only if w is securable with respect to R and a. Then we can take $\xi(a) = \xi(a, 1)$.

We choose the function $\xi(a, w)$ to satisfy

$$(25.1) \qquad \xi(a, w) = \begin{cases} 0 \text{ if } \overline{\mathrm{Seq}}(w) \text{ (Case 0)}, \\ 1 \text{ if } w \text{ is secured (Case 1)}, \\ 3 \cdot 5^{f_{a,w}} \text{ if } \mathrm{Seq}(w) \text{ but } w \text{ is not secured (Case 2)}, \end{cases}$$

where (in Case 2) $f_{a,w}$ defines recursively $\sum_{s<n} {}_O (\xi(a, w * 2^{s+1}) +_O 1_O)$ as a function of n_O.

---

[28] Our proof of the implication $\xi(a) \varepsilon O \rightarrow (\alpha)(Ex)R(a, \alpha, x)$ makes use of the classical second number class.—Putting $\xi_1(a) = \xi(a) +_O a_O$, ξ_1 has the properties of ξ in the theorem, and $a \neq b \rightarrow \xi_1(a) \neq \xi_1(b)$; so the decision problem for $\hat{a}(\alpha)(Ex)R(a, \alpha, x)$ is 1-1 reducible (Post) to that for O (cf. [19], [11] p. 343 Example 2).

We obtain such a function ξ as follows. Let

$$\theta(z,a,w,b) \cong \sum\nolimits_{O\ s<\mathrm{nat}(b)} (\Phi_2(z,a,w{*}2^{s+1}) +_O 1_O)$$

(cf. **21**). Let p define θ recursively. Let

$$(25.2)\quad \psi(z,a,w) = \begin{cases} 0 \text{ if } \overline{\mathrm{Seq}}(w), \\ 1 \text{ if } \mathrm{Seq}(w) \;\&\; (Et)_{t\leq \mathrm{lh}(w)} R(a, \prod_{i<t} p_i^{(w)_i}), \\ 3\cdot 5^{S_1^3(p,z,a,w)} \text{ if } \mathrm{Seq}(w) \;\&\; \overline{(Et)}_{t\leq \mathrm{lh}(w)} R(a, \prod_{i<t} p_i^{(w)_i}). \end{cases}$$

Then ψ is primitive recursive. Choose e by the recursion theorem so that e defines $\psi(e,a,w)$ recursively.[20] Let $\xi(a,w) = \psi(e,a,w)$.

Now we prove by an induction of form corresponding to the inductive definition $S1$-$S3$ of securability that, if w is securable, then $\xi(a,w) \,\varepsilon\, O$.

1. Assume w is secured. Then by Case 1 of (25.1), $\xi(a,w) = 1$. So by $O1$, $\xi(a,w) \,\varepsilon\, O$.

2. Suppose Seq(w), w is not secured, but, for every s, $w{*}2^{s+1}$ is securable, and as hypothesis of the induction that, for every s, $\xi(a,w{*}2^{s+1}) \,\varepsilon\, O$. Then by (XVI) and (XXI), for each n, $\sum_{O\ s<n} (\xi(a,w{*}2^{s+1}) +_O 1_O) \,\varepsilon\, O$. Moreover by (XXIII) and (XVIII), for each n,

$$\sum\nolimits_{O\ s<n} (\xi(a,w{*}2^{s+1}) +_O 1_O) <_O \sum\nolimits_{O\ s<n+1} (\xi(a,w{*}2^{s+1}) +_O 1_O).$$

Hence by (25.1) Case 2 with $O3$, $\xi(a,w) \,\varepsilon\, O$.

Conversely, we prove that, if $\xi(a,w) = c$ and $c \,\varepsilon\, O$, then w is securable. We give the proof by transfinite induction on $|c|$.[29] The cases in the proof correspond to the clauses $O1$-$O3$ by which c can $\varepsilon\, O$.

1. Assume $c = 1$. Then $\xi(a,w) = 1$; it is Case 1 of (25.1) which applies to $\xi(a,w)$; and so w is secured, a fortiori securable.

2. Assume $c = 2^y$ where $y \,\varepsilon\, O$, so by (I), $y \neq 0$. But then $\xi(a,w) = c$ is impossible under (25.1).

---

[29] Several attempts to get by with induction on c instead of on $|c|$ were unsuccessful. A difficulty is that $a +_O b$ is not necessarily $\geqq_O b$. Indeed there is no function $\circ$ such that, for all $a, b \,\varepsilon\, O$, both $a \circ b \geqq_O a$ and $a \circ b \geqq_O b$. For if e and f define recursively n_O and $(n+1)_O$, respectively, as functions of n_O, and $a = 3\cdot 5^e$ and $b = 3 \cdot 5^f$, then $|a| = |b| = \omega$, but $a \neq b$; so by the linearity of the ordering by $<_O$ of the members of O which are $<_O$ a given one and the consistency of this ordering with that of the ordinals represented (end **20**), a and b cannot both occur among the numbers $<_O 2^{a\circ b}$. Another difficulty is that the associative law $(a +_O b) +_O c = a +_O (b +_O c)$ fails, e.g., when $a = 1$, $b = 2$, $c = 3\cdot 5^y$.

3. Assume $c = 3 \cdot 5^y$ where, for each n, $\Phi(y, n_O)$ is defined (call it y_n), $y_n \,\varepsilon\, O$ and $y_n <_O y_{n+1}$ (and by $O3$, $y_n <_O c$). Now it is Case 2 of (25.1) which applies to $\xi(a, w)$. Thus Seq(w), but w is not secured; moreover (as $y = f_{a,w}$), for each n, $y_n = \underset{s<n}{\Sigma_O} (\xi(a, w{*}2^{s+1}) +_O 1_O)$, so

$$\underset{s<n}{\Sigma_O} (\xi(a, w{*}2^{s+1}) +_O 1_O) \,\varepsilon\, O \text{ and } \underset{s<n}{\Sigma_O} (\xi(a, w{*}2^{s+1}) +_O 1_O) <_O c.$$

Using (XXII) and (XX), for every s, $\xi(a, w{*}2^{s+1}) \,\varepsilon\, O$. Furthermore, since (taking n above $= s + 1$)

$$\begin{aligned} c >_O \underset{m<s+1}{\Sigma_O} (\xi(a, w{*}2^{m+1}) +_O 1_O) \\ = \underset{m<s}{\Sigma_O} (\xi(a, w{*}2^{m+1}) +_O 1_O) +_O (\xi(a, w{*}2^{s+1}) +_O 1_O), \end{aligned}$$

$$|\,\xi(a, w{*}2^{s+1})\,| < |\,\xi(a, w{*}2^{s+1}) +_O 1_O\,| \leqq |\, \underset{m<s+1}{\Sigma_O} (\xi(a, w{*}2^{m+1}) +_O 1_O\,| < |\,c\,|.$$

So by the hypothesis of the transfinite induction, since $\xi(a, w{*}2^{s+1}) \,\varepsilon\, O$, $w{*}2^{s+1}$ is securable. This is for every s, and Seq(w), so w is securable.

COROLLARY. *The predicates $a \,\varepsilon\, O$ and $a <_O b$ are not expressible in the forms $(E\alpha)(x)R(a, \alpha, x)$ and $(E\alpha)(x)S(a, b, \alpha, x)$ with general recursive R and S, respectively; a fortiori, they are not arithmetical.*

Proof. Were $a \,\varepsilon\, O \equiv (E\alpha)(x)R(a, \alpha, x)$ with R recursive, by applying the theorem to $T_1{}^{\alpha}(a, a, x)$ as its $R(a, \alpha, x)$, we would have

$$(\alpha)(Ex)T_1{}^{\alpha}(a, a, x) \equiv (E\alpha)(x)R(\xi(a), \alpha, x) \equiv (E\alpha)(x)R_1(a, \alpha, x)$$

with recursive R_1, which is impossible.[30] The statement for $a <_O b$ follows, since $a \,\varepsilon\, O \equiv a <_O 2^a$ by $O2$ and (IV).

26. A digression concerning recursive linear orderings. Let us examine further the construction used in the proof of Theorem I.

We use $\succ$ to denote the linear ordering of the set $\hat{w}\,$Seq(w) of the numbers w of the form $\bar{\alpha}(x)$ which is established by ordering the finite sequences $\alpha(0), \cdots, \alpha(x-1)$ lexicographically with the infinite alphabet $\cdots, 2, 1, 0$. In this ordering, 1 (which represents the empty sequence) is the highest element. If $\overline{\text{Seq}}(u)$ or $\overline{\text{Seq}}(v)$, $u \succ v$ shall be false. Then the predicate $u \succ v$ is primitive recursive.

Now, for any given predicate $R(a, v)$ and numbers a and w, we define a set $S^{R,a}{}_w$, or briefly $S^a{}_w$ or S_w, as follows. If w is secured (with respect to R and a), S_w shall be the unit set $\{w\}$. If Seq(w) and w is unsecured, S_w

---

[30] [12] Theorem 3 Proof for $n > 0$. Also cf. [12] Theorem 5.

shall consist of all $w*\bar{\alpha}(t)$ which are unsecured or immediately secured. If $\overline{\mathrm{Seq}}(w)$, S_w shall be the empty set. Then $u \varepsilon S^{R,a}{}_w$ as predicate of u, w, a is primitive recursive uniformly in R.

If $u \varepsilon S_w$, then $\mathrm{Seq}(u)$. So if $\mathrm{Seq}(w)$, S_w is linearly ordered by $\succ$ with w is highest element.

(F) *The set* S_w *is well-ordered by* $\succ$, *if* w *is securable, in which case the ordinal number of* S_w *is* $|\xi(a, w)| + 1$.[31]

We prove this, as follows, by an induction of form corresponding to the inductive definition of securability.

1. Assume w is secured. Then S_w is $\{w\}$, which is well-ordered with ordinal number 1. But by Case 1 of (25.1), $\xi(a, w) = 1 = 0_O$. So $|\xi(a, w)| + 1 = 0 + 1 = 1$.

2. Assume $\mathrm{Seq}(w)$, w is unsecured, but $w*2^{s+1}$ is securable $(s = 0, 1, 2, \cdots)$, and as hypothesis of the induction that $S_{w*2^{s+1}}$ is well-ordered by $\succ$ with ordinal number $|\xi(a, w*2^{s+1})| + 1$. Then under the ordering $\succ$, S_w is the sum of $\{w\}, \cdots, S_{w*2^3}, S_{w*2^2}, S_{w*2^1}$, and so is well-ordered with the ordinal number $\sum_{s=0,1,2,\cdots} (|\xi(a, w*2^{s+1})| + 1) + 1$. But by Case 2 of (25.1),

$$\sum_{s=0,1,2,\cdots} (|\xi(a, w*2^{s+1})| + 1) + 1 = |\xi(a, w)| + 1.$$

(G) *If* $\mathrm{Seq}(w)$, *and* S_w *is well-ordered by* $\succ$, *then* w *is securable.*

For we can show by transfinite induction on the ordinal number of S_w that, if $\mathrm{Seq}(w)$, and S_w is well-ordered, then each member $w*\bar{\alpha}(t)$ of S_w is securable; so in particular, then w $(= w*\bar{\alpha}(0))$ is securable.

Now we return to (F), and give another argument for the part not concerning ξ, in which we show that, if S_w were not well-ordered by $\succ$, w would not be securable.

Accordingly suppose S_w is not well-ordered by $\succ$, so there there is an infinite descending sequence $\beta(0) \succ \beta(1) \succ \beta(2) \succ \cdots$ within S_w. Of course then $\mathrm{Seq}(w)$ and w is unsecured. Write $w = \bar{\gamma}(x)$ (then $x = \mathrm{lh}(w)$).

Now we shall infer by induction on t that, for all sufficiently great m (say $m \geqq m_t$), $\mathrm{lh}(\beta(m)) > x + t$ and $(\beta(m))_x \dot{-} 1, \cdots, (\beta(m))_{x+t} \dot{-} 1$ are fixed in value (call their values then $\alpha(0), \cdots, \alpha(t)$, respectively). Let

---

[31] That S_w is well-ordered by $\succ$ if w is securable was argued in substance by Brouwer [1] Section 1 or [2] Section 2. For w securable, S_w consists of w and, if w is unsecured, those numbers $w*\bar{\alpha}(t)$ coming below w in the lexicographic ordering $\succ$ which under the inductive definition of securability must be recognized as securable before w itself can be so recognized.

$n_t = m_{t-1}$ if $t > 0$ $(= 0$ if $t = 0)$, so that, using the hypothesis of the induction, for all $m \geqq n_t$, $\beta(m)$ represents a sequence the first $x + t$ values of which are fixed as $\gamma(0), \cdots, \gamma(x-1), \alpha(0), \cdots, \alpha(t-1)$. Then by the lexicographic ordering, for $m > n_t$, $\beta(m)$ must represent a sequence $\gamma(0), \cdots, \gamma(x-1), \alpha(0), \cdots, \alpha(t-1), (\beta(m))_{x+t} \dot{-} 1, \cdots, (\beta(m))_{\mathrm{lh}(\beta(m))-1} \dot{-} 1$ of length $> x + t$, and with increasing m $(m > n_t)$ $(\beta(m))_{x+t} \dot{-} 1$ is monotone non-increasing. But since $(\beta(m))_{x+t} \dot{-} 1$ is a natural number, starting with its value for $m = n_t + 1$, it can decrease only a finite number of times, so beginning with a certain value m_t $(> n_t)$ of m, $(\beta(m))_{x+t} \dot{-} 1$ remains fixed, call its value then $\alpha(t)$.

We can express α in terms of β and w by the formula[32]

$$\alpha(t) = (\epsilon s(m)_{m \geqq (s)_0}[(\beta(m))_{\mathrm{lh}(w)+t} \dot{-} 1 = (s)_1])_1.$$

So α is arithmetical uniformly in β.[33]

Since $\beta(m_t)$ is of the form $w * \bar{\alpha}(t+1) * v$, and $\beta(m_t) \varepsilon S_w$, $w * \bar{\alpha}(t)$ and w are unsecured; a fortiori, $\bar{R}(a, w * \bar{\alpha}(t))$.

Thus $(E\alpha)(t)\bar{R}(a, w * \bar{\alpha}(t))$, whence $\overline{(\alpha)}(Et)R(a, w * \bar{\alpha}(t))$; besides which w is unsecured; so w is unsecurable.

So we have another proof of (F). In the course of this we have shown:

(H) *For each w, there is a function α arithmetical uniformly in β such that, if $\beta(0), \beta(1), \beta(2), \cdots \varepsilon S_w$ and $\beta(0) \succ \beta(1) \succ \beta(2) \succ \cdots$, then w is unsecured and $(t)\bar{R}(a, w * \bar{\alpha}(t))$.*

These results, of course, hold with $a_1, \cdots, a_n$ (for any fixed $n \geqq 0$) in place of a. Now we give an application with $w = 1$ and $n = 0$.

Elsewhere we have shown that there is a primitive recursive predicate $R(v)$ such that (a) $(E\alpha)(x)\bar{R}(\bar{\alpha}(x))$, but (b) for no arithmetical α, $(x)\bar{R}(\bar{\alpha}(x))$.[34] Consider the set S^{R}_1 or briefly S_1. By (a), $\overline{(\alpha)}(Ex)R(\bar{\alpha}(x))$, i.e., 1 is not securable, so by (G), S_1 is not well-ordered by $\succ$, so there is a β such that $\beta(0), \beta(1), \beta(2), \cdots \varepsilon S_1$ and $\beta(0) \succ \beta(1) \quad \beta(2) \succ \cdots$. But were there such a β which was arithmetical, the corresponding α of (H) would be an arithmetical function such that $(x)\bar{R}(\bar{\alpha}(x))$, contradicting (b).

The set S_1 is primitive recursive and infinite, and $0 \not\varepsilon S_1$ but $1 \varepsilon S_1$.

---

[32] As in [7] or [11] p. 317 (following Gödel) $\epsilon s A(s) = \{$the least s such that $A(s)\}$ if $(Es)A(s)$, $= 0$ otherwise.

[33] [11] pp. 239, 291, 292, or [12] Section 2. A function α is arithmetical (arithmetical in β), if its representing predicate $\alpha(x) = w$ is arithmetical (arithmetical in β). Cf. [11] pp. 285, 318.

[34] [12] Corollary to Theorem 7. We are rewriting the $(x)R(\alpha, x)$ there as $(x)\bar{R}(\bar{\alpha}(x))$ (cf. beginning **24** above and [11] p. 228 #D).

Let $\sigma(0) = 1$, $\sigma(n+1) = n+1$ if $n+1 \varepsilon S_1$, and $\sigma(n+1) = \sigma(n)$ if $n+1 \not\varepsilon S_1$. Then σ is primitive recursive, and $\sigma(0), \sigma(1), \sigma(2), \cdots$ is an enumeration of S_1 with only consecutive repetitions. Now define

$$i >^R j \equiv [\sigma(i) \neq \sigma(j) \mathbin{\&} \sigma(i) \succ \sigma(j)] \vee [\sigma(i) = \sigma(j) \mathbin{\&} i > j].$$

Then $i >^R j$ is a primitive recursive linear ordering of the natural numbers, and members of S_1 are ordered by $>^R$ as they are by $\succ$.

So for any infinite descending sequence $\beta(0) \succ \beta(1) \succ \beta(2) \succ \cdots$ within S_1, $\beta(0) >^R \beta(1) >^R \beta(2) >^R \cdots$. Thus there is an infinite descending sequence in the ordering $>^R$ of the natural numbers.

But there is none which is arithmetical. For given any γ such that $\gamma(0) >^R \gamma(1) >^R \gamma(2) >^R \cdots$, let $\psi(0) = 0$ and

$$\psi(n+1) = \mu t_{t > \psi(n)} [\sigma(\gamma(t)) \neq \sigma(\gamma(\psi(n)))],$$

and let $\beta(n) = \sigma(\gamma(\psi(n)))$. Then β is general recursive, a fortiori arithmetical, in γ, and $\beta(0) \succ \beta(1) \succ \beta(2) \succ \cdots$ is an infinite descending sequence within S_1. Thus we have established:

(J) *There is a primitive recursive linear ordering* $>^R$ *of the natural numbers such that, for some number-theoretic function* γ,

$$\gamma(0) >^R \gamma(1) >^R \gamma(2) >^R \cdots,$$

but, for every arithmetical γ, $(En)[\gamma(n) \leqq^R \gamma(n+1)]$.

In other words, $>^R$ is a well-ordering of the natural numbers with respect to arithmetical sequences $\gamma(0), \gamma(1), \gamma(2), \cdots$, but not with respect to arbitrary sequences.[35]

In fact, the primitive recursive $R(v)$ can be chosen so that $>^R$ is a well-ordering with respect to a much larger class C of sequences $\gamma(0), \gamma(1), \gamma(2), \cdots$ than the arithmetical sequences, but not with respect to arbitrary sequences.[36]

---

[35] G. Kreisel raised (in a different formulation) the question whether this is the case in a letter to the author dated December 9, 1952 (in reply to a letter from the author dated November 13, 1952, in which the result of [12] which leads here to (J) was stated), and also in a communication received by Leon Henkin November 20, 1952. Kreisel formulated the question in relation to the formal system Z_μ of Hilbert-Bernays [5] and some ω-consistent formal system S of analysis including Z_μ, while here we have considered the matter only informally. The arithmetical functions are however exactly those expressible by terms of Z_μ under the usual interpretation of the symbolism, and our arguments about arbitrary number-theoretic functions can presumably be formalized in a formal system S of analysis of a usual sort.

[36] We need only start from Theorems 7 and 9 of [12] (instead of from Corollary

27. Reduction of $a \varepsilon Q$ (amended). Now we begin the proof that $a \varepsilon O \equiv (\alpha)(Ey)R(a, \alpha, y)$ for a suitable primitive recursive R. As before, we start from

$$a \varepsilon O \equiv a \varepsilon Q \tag{28}$$

(end **12**) and (32) (in **13**, **20**). But now we amend the former treatment (cf. **14**) so that we shall be able to give (in **28**) the necessary argument to supplement the application of the reduction technique (cf. **18**).

The inductive definition of Q (in **12**) can be restated as follows (cf. beginning **14**[11]).

$Q_1 1$. If $a = 1$, then $a \varepsilon Q$. $Q_1 2$. If $a = 2^{(a)_0}$ & $(a)_0 \varepsilon Q$, then $a \varepsilon Q$. $Q_1 3$. If

$$\begin{aligned} a = 3 \cdot 5^{(a)_2} \,\&\, (x_0)(Ey_0)T((a)_2, (x_0)_O, y_0) \\ \&\, (x_0)(x_1)[T((a)_2, (x_0)_O, x_1) \rightarrow U(x_1) \varepsilon Q] \\ \&\, (x_0)(x_1)(x_2)[T((a)_2, (x_0)_O, x_1) \\ \&\, T((a)_2, (x_0 + 1)_O, x_2) \rightarrow (Ey_1)V(U(x_1), U(x_2), y_1)], \end{aligned}$$

then $a \varepsilon Q$. $Q_1 4$. $a \varepsilon Q$ only as required by $Q_1 1$-$Q_1 3$.

By writing a corresponding equivalence (as we did for various inductive definitions in the first paper; cf. **18** here), and then substituting in it for "$a \varepsilon Q$" the predicate expression "$(\alpha)(Ey)R(a, \bar{\alpha}(y))$" where $R(a, v)$ remains to be selected, we obtain

$$\begin{aligned} (\alpha)(Ey)R(a, \bar{\alpha}(y)) \equiv a = 1 \vee \{a = 2^{(a)_0} \,\&\, (\alpha)(Ey)R((a)_0, \bar{\alpha}(y))\} \\ \vee \{a = 3 \cdot 5^{(a)_2} \,\&\, (x_0)(Ey_0)T((a)_2, (x_0)_O, y_0) \,\&\, (x_0)(x_1)[T((a)_2, (x_0)_O, x_1) \\ \rightarrow (\alpha)(Ey)R(U(x_1), \bar{\alpha}(y))] \,\&\, (x_0)(x_1)(x_2)[T((a)_2, (x_0)_O, x_1) \\ \&\, T((a)_2, (x_0 + 1)_O, x_2) \rightarrow (Ey_1)V(U(x_1), U(x_2), y_1)]\}. \end{aligned} \tag{27.1}$$

We shall presently define $R(a, v)$ so that it will be primitive recursive and (27.1) will be true. When this has been done, it will follow from (27.1) by an induction of form corresponding to the inductive definition $Q_1 1$-$Q_1 4$ (cf. **18**) that

$$a \varepsilon Q \rightarrow (\alpha)(Ey)R(a, \bar{\alpha}(y)). \tag{27.2}$$

---

to Theorem 7) to obtain as C, for any fixed $y \varepsilon O$, the sequences recursive in the predicate H_y of [12] Section 6; we apply Theorem 7 to $H_{y +_O 2_O}$ to allow for the definition of α from β for (H). In [13] XXVI we obtain (with the aid of Theorem II below) an improvement over [12] Theorem 7, using which C can be the sequences recursive in H_y for any $y \varepsilon O$.

Note that, if C and D are independent of α,

$$\bar{C} \vee \bar{D} \rightarrow \{(\alpha)(C \,\&\, A(\alpha)) \vee (\alpha)(D \,\&\, B(\alpha)) \equiv (\alpha)[(C \,\&\, A(\alpha)) \vee (D \,\&\, B(\alpha))]\}. \tag{27.3}$$

To find a primitive recursive $R(a, v)$ satisfying (27.1), we begin by observing that in the right member of (27.1) (call it (a)) the quantifiers can be advanced (using (27.3) to bring both (α)'s across the second $\vee$ as one (α))[37] to give an equivalent of the form

$$(x_0)(x_1)(x_2)(\alpha)(Ey_0)(Ey_1)(Ey)M^R(a, x_0, x_1, x_2, \bar{\alpha}(y), y_0, y_1) \tag{b}$$

where $M^R(a, x_0, x_1, x_2, \bar{\alpha}(y), y_0, y_1)$ is exactly (a) with its quantifiers omitted. Next (b) is equivalent to[38]

$$(\alpha)(Ey)\{M^R(a, \alpha(0), \alpha(1), \alpha(2), \overline{\{\lambda i \alpha(i+3)\}}((y \dot{-} 3)_2), (y \dot{-} 3)_0, (y \dot{-} 3)_1) \,\&\, y \geqq 3\}. \tag{c}$$

But $y = \mathrm{lh}(\bar{\alpha}(y))$, and for $y \geqq 3$,

$$\alpha(0) = (\bar{\alpha}(y))_0 \dot{-} 1, \alpha(1) = (\bar{\alpha}(y))_1 \dot{-} 1, \alpha(2) = (\bar{\alpha}(y))_2 \dot{-} 1$$

and (since $y \geqq 3 \,\&\, i < (y \dot{-} 3)_2 \rightarrow i + 3 < y$)

$$\overline{\{\lambda i \alpha(i+3)\}}((y \dot{-} 3)_2) = \prod_{i < (y \dot{-} 3)_2} p_i \exp(\bar{\alpha}(y))_{i+3},$$

using which (c) takes the form $(\alpha)(Ey)N^R(a, \bar{\alpha}(y))$ (call it (d)). Now we propose to define R by

$$R(a, v) \equiv N^R(a, v). \tag{27.4}$$

This does define a predicate $R(a, v)$ by course-of-values recursion on v, since for $\mathrm{lh}(v) < 3$ it makes $R(a, v)$ false outright, while for $\mathrm{lh}(v) \geqq 3$ it makes $R(a, v)$ depend primitive recursively on $R(b, w)$ only for

$$w = \prod_{i < (\mathrm{lh}(v) \dot{-} 3)_2} p_i \exp(v)_{i+3} < v.$$

The predicate $R(a, v)$ thus defined is primitive recursive.[39] Substituting into (27.4) and quantifying,

$$(\alpha)(Ey)R(a, \bar{\alpha}(y)) \equiv (\alpha)(Ey)N^R(a, \bar{\alpha}(y)), \tag{27.5}$$

---

[37] See **14** p. 54 for justification of the advances intuitionistically.

[38] By [12] Footnote 7, and a modification of **5** (16),[11] [9] (15) or [11] p. 285 (17). Following Church, $\lambda i f(i)$ is $f(i)$ as a function of i (cf. [11] p. 34).

[39] One can eliminate both the course-of-values character of the recursion on v and the substitution for the parameter a by the methods of Péter [17] Sections 1 and 2 or [18] Sections 3 and 5. Actually by beginning **24** it would suffice for our results to observe merely that $R(a, v)$ is general recursive, which is obvious.

whence by reversing the steps from (a) to (d) we now prove (27.1) and thence (27.2) for the R defined by (27.4).

28. Supplementary argument to the reduction of $a \varepsilon Q$. We prove that

$$(28.1) \qquad (\alpha)(Ey)R(a, \bar{\alpha}(y)) \rightarrow a \varepsilon Q,$$

conversely to (27.2). So assume for a given a that $(\alpha)(Ey)R(a, \bar{\alpha}(y))$ for the R defined by (27.4). Then by **24**, 1 is securable with respect to (R and) a; and by **26** (F), the set $S^{R,a}{}_1$ or briefly $S^a{}_1$ is well-ordered by $\succ$, say with the ordinal number $|S^a{}_1|$ $(= |\xi(a, 1)| + 1)$. We use transfinite induction on this number.[40] By our assumption and (27.1), one of three cases applies.

1. $a = 1$. Then $a \varepsilon Q$, by $Q_1 1$.

2. $a = 2^{(a)_0} \,\&\, (\alpha)(Ey)R((a)_0, \bar{\alpha}(y))$. By **24**, 1 is securable with respect to $(a)_0$. For any x_0, x_1, x_2, α, let y^α be the least y such that $R((a)_0, \bar{\alpha}(y))$, so $\bar{\alpha}(y^\alpha)$ is immediately secured with respect to $(a)_0$; and consider some function β with $\beta(3 + y^\alpha) = [x_0, x_1, x_2] * \bar{\alpha}(y^\alpha)$, where $[x_0, \cdots, x_n]$ abbreviates $p_0{}^{x_0+1} \cdots p_n{}^{x_n+1}$. For $z < 3$, $\bar{R}(a, \beta(z))$ immediately by the definition of $R(a, v)$, and for $3 \leqq z < 3 + y^\alpha$ because $R(a, \bar{\beta}(z))$ would imply $R((a)_0, \bar{\alpha}((z \dot{-} 3)_2))$ with $(z \dot{-} 3)_2 < y^\alpha$ contradicting the choice of y^α; but we can choose $z \geqq 3 + y^\alpha$ so that $(z \dot{-} 3)_2 = y^\alpha$, and then $R(a, \bar{\beta}(z))$. Thus there is a u such that $[x_0, x_1, x_2] * \bar{\alpha}(y^\alpha) * u$ is immediately secured with respect to a (namely $\bar{\beta}(\mu z R(a, \bar{\beta}(z)))/\bar{\beta}(3 + y^\alpha)$). So $S^a{}_{[x_0, x_1, x_2]}$ includes all the numbers $[x_0, x_1, x_2] * w$ for $w \varepsilon S^{(a)_0}{}_1$, hence $|S^a{}_{[x_0, x_1, x_2]}| \geqq |S^{(a)_0}{}_1|$. Moreover, under the ordering $\succ$, $S^a{}_{[x_0, x_1]}$ is the sum of $\{[x_0, x_1]\}$ and the sets $S^a{}_{[x_0, x_1, x_2]}$ for $x_2 = \cdots, 2, 1, 0$, hence $|S^a{}_{[x_0, x_1]}| \geqq \omega |S^{(a)_0}{}_1| + 1$; etc. So finally $|S^a{}_1| \geqq \omega(\omega(\omega |S^{(a)_0}{}_1| + 1) + 1) + 1 > |S^{(a)_0}{}_1|$. So by the hypothesis of the induction, $(a)_0 \varepsilon Q$. By $Q_1 2$, $a \varepsilon Q$.

3. $a = 3 \cdot 5^{(a)_2} \,\&\, (x_0)(Ey_0)T((a)_2, (x_0)_O, y_0) \,\&\, (x_0)(x_1)[T((a)_2, (x_0)_O, x_1)$
$\rightarrow (\alpha)(Ey)R(U(x_1), \bar{\alpha}(y))] \,\&\, (x_0)(x_1)(x_2)[T((a)_2, (x_0)_O, x_1)$
$\& \, T((a)_2, (x_0 + 1)_O, x_2) \rightarrow (Ey_1)V(U(x_1), U(x_2), y_1)]$.

In particular, for each x_0 and x_1 such that $T((a)_2, (x_0)_O, x_1)$, 1 is securable with respect to $U(x_1)$. For any such x_0, x_1 fixed for the moment, and any x_2, α, let y^α be the least y such that $R(U(x_1), \bar{\alpha}(y))$, so $\bar{\alpha}(y^\alpha)$ is immediately secured with respect to $U(x_1)$; and consider some function β with $\bar{\beta}(3+y^\alpha) = [x_0, x_1, x_2] * \bar{\alpha}(y^\alpha)$. For $z < 3$, $\bar{R}(a, \bar{\beta}(z))$ obviously, and for $3 \leqq z < 3 + y^\alpha$ because $R(a, \bar{\beta}(z))$ would imply $R(U(x_1), \bar{\alpha}((z \dot{-} 3)_2))$ with $(z \dot{-} 3)_2 < y^\alpha$ contradicting the choice of y^α; but we can choose $z \geqq 3 + y^\alpha$ so that $(z \dot{-} 3)_2 = y^\alpha$, $T((a)_2, (x_0)_O, (z \dot{-} 3)_0)$ and

---

[40] There is a longer proof, which avoids the use of the classical second number class, and uses instead (E) and an induction corresponding to the inductive definition of securability.

$$T((a)_2, (x_0+1)_O, x_2) \rightarrow V(U(x_1), U(x_2), (z \dot{-} 3)_1),$$

and then $R(a, \bar{\beta}(z))$. Thus there is a u such that $[x_0, x_1, x_2] * \bar{\alpha}(y^\alpha) * u$ is immediately secured with respect to a. Hence $|S^a{}_1| \geqq \omega \, |S^{U(x_1)}{}_1| + 3 > |S^{U(x_1)}{}_1|$. So applying the hypothesis of the induction,

$$(x_0)(x_1)[T((a)_2, (x_0)_O, x_1) \rightarrow U(x_1) \varepsilon Q].$$

By $Q_1 3$, $a \varepsilon Q$.[41]

29. Classification of $a \varepsilon O$ and $a <_O b$ from above.

THEOREM II. *For certain primitive recursive predicates R and S, $a \varepsilon O \equiv (\alpha)(Ey)R(a, \alpha, y)$ and $a <_O b \equiv (\alpha)(Ey)S(a, b, \alpha, y)$.*

Proof. Letting $R(a, \alpha, y) \equiv R(a, \bar{\alpha}(y))$ for the $R(a, v)$ defined by (27.4), and combining (28) (**12** or **27**) with (27.2) and (28.1),

$$a \varepsilon O \equiv (\alpha)(Ey)R(a, \alpha, y). \tag{29.1}$$

By (IV) (**12** or **20**) with (21.1),

$$a <_O b \equiv b \varepsilon O \,\&\, (Ex)V(a, b, x), \tag{29.2}$$

which with (29.1) by advancing and contracting quantifiers gives

$$a <_O b \equiv (\alpha)(Ey)[R(b, \alpha, (y)_0) \,\&\, V(a, b, (y)_1)]. \tag{29.3}$$

30. The predicates $a \varepsilon O^P$ and $a <_O{}^P b$. The definitions of the predicates $a \varepsilon O$ and $a <_O b$ and the representation of ordinals can be relativized by using recursiveness in a given predicate $P(a)$ in place of recursiveness.[42] To do this, we simply read throughout the first three and last paragraphs of **20** "T^P", "$T_1{}^P$", "Φ^P", "$\Phi_1{}^P$", "recursively from P", "O^P", "$<_O{}^P$", "$|\ |^P$", "$\omega_1{}^P$" for "T", "T_1", "Φ", "Φ_1", "recursively", "O", "$<_O$", "$|\ |$", "ω_1", respectively.[43] The resulting notions O^P, $<_O{}^P$, $|\ |^P$, $\omega_1{}^P$ (e.g., for $P(a) \equiv a \varepsilon O$) are discussed in [12] 6.4 ff. and [13] Section 7 (also cf. [14] end 3.5).

---

[41] By this proof and **27**, $a \varepsilon O \rightarrow |S^a{}_1| \geqq |a|$.

[42] Or indeed in any given list Ψ of l functions and predicates of $m_1, \cdots, m_l$ variables, respectively.

[43] For $T_1{}^P$, $\Phi_1{}^P$, or what is the same $T_1{}^\pi$, $\Phi_1{}^\pi$ where π is the representing function of P, cf. [11] pp. 227, 292, 341.

The above arguments carry over to these relativized notions. For the most part, the changes required in the text are only the supplying of superscripts "^{P}" on the symbols for various functions, predicates and sets, and the reading of "recursive in P" (which can often be "recursive uniformly in P") in place of "recursive".

Thus in **20** and **21**, then T^P, V^P, enm^P and $m <_b'^P n$ are primitive, and Φ^P and $m <_b^P n$ partial, recursive uniformly in P (but n_O, $\text{Fin}(b)$ and $\text{nat}(b)$ do not change). In **22** and **23** we have a function $+_O{}^1$ and operation $\sum_O{}^1$,[44] which are still primitive recursive (in Case 3 for (22.1), $d_{a,y}$ defines recursively $a +_O{}^1 \Phi^P(y, n_O)$ as a function of n_O uniformly from P; $\theta^P(z, a, y, n) \cong \Phi_2(z, a, \Phi^P(y, n))$; p^1 is a uniform Gödel number from P of $\theta^P(z, a, y, n)$; in (22.2) $S_1{}^3$ becomes $S_1{}^{3,1}$). In **24**, using $(Ex)R^P(a, \alpha, x) \equiv (Ex)T_1{}^{1,1}(\bar{\pi}(x), \bar{\alpha}(x), f, a)$ for π the representing function of P, any predicate of the form $(Ex)R^P(a, \alpha, x)$ with R^P general recursive in P is expressible as $(Ex)R^P(a, \bar{\alpha}(x))$ with an R^P primitive recursive in P (uniformly so, if $R^P(a, \alpha, x)$ was); and the definitions of 'secured' etc. are considered with respect to R^P and a (but $\text{Seq}(w)$ does not change). In **25** Theorem I becomes: *Given any predicate $R^P(a, \alpha, x)$ general recursive in P (uniformly in P), a function $\xi^P(a)$ primitive recursive in P (uniformly in P) can be found such that $(\alpha)(Ex)R^P(a, \alpha, x) \equiv \xi^P(a) \varepsilon O^P$.* In **26** $\succ$ is unchanged; and replacing R by R^P in our arguments, (J) gives a linear ordering $>^{R^P}$, primitive recursive in P, which is a well-ordering with respect to sequences $\gamma(0), \gamma(1), \gamma(2), \cdots$ arithmetical in P, but not with respect to arbitrary sequences. In **27-29**, starting from $a \varepsilon O^P \equiv a \varepsilon Q^P$, we obtain as the relative version of Theorem II: *For certain predicates R^P and S^P primitive recursive uniformly in P,*

$$a \varepsilon O^P \equiv (\alpha)(Ey)R^P(a, \alpha, y) \quad \textit{and} \quad a <_O{}^P b \equiv (\alpha)(Ey)S^P(a, b, \alpha, y).$$

THE UNIVERSITY OF WISCONSIN.

---

[44] The superscript "^{1}" is for the number 1 of the arguments of P; were we relativizing to Ψ we would have $+_O{}^{m_1, \cdots, m_l}$, $\sum_O{}^{m_1, \cdots, m_l}$.[42]

BIBLIOGRAPHY.

[1] L. E. J. Brouwer, "Beweis, dass jede volle Funktion gleichmässig stetig ist," *Koninklijke Nederlandsche Akademie van Wetenschappen, Proceedings of the Section of Sciences*, vol. 27 (1924), pp. 189-193.

[2] ———, "Über Definitionsbereiche von Funktionen," *Mathematische Annalen*, vol. 97 (1927), pp. 60-75.

[3] A. Church, "The constructive second number class," *Bulletin of the American Mathematical Society*, vol. 44 (1938), pp. 224-232.

[4] A. Church and S. C. Kleene, "Formal definitions in the theory of ordinal numbers," *Fundamenta Mathematicae*, vol. 28 (1936), pp. 11-21.

[5] D. Hilbert and P. Bernays, *Grundlagen der Mathematik*, Berlin, vol. 2 (1939).

[6] L. Kalmár, "Egyszerü példa eldönthetetlen aritmetikai problémára (Ein einfaches Beispiel für ein unentscheidbares arithmetisches Problem)," *Matematikai és Fizikai Lapok*, vol. 50 (1943), pp. 1-23.

[7] S. C. Kleene, "General recursive functions of natural numbers," *Mathematische Annalen*, vol 112 (1936), pp. 727-742.

[8] ———, "On notation for ordinal numbers," *Journal of Symbolic Logic*, vol. 3 (1938), pp. 150-155.

[9] ———, "Recursive predicates and quantifiers," *Transactions of the American Mathematical Society*, vol. 53 (1943), pp. 41-73 (cf. **19** above).

[10] ———, "Recursive functions and intuitionistic mathematics," *Proceedings of the International Congress of Mathematicians (Cambridge, Mass., U. S. A., Aug. 30-Sept. 6, 1950)*, 1952, vol. 1, pp. 679-685.

[11] ———, *Introduction to Metamathematics*, New York, Toronto, Amsterdam and Groningen (1952).[23]

[12] ———, "Arithmetical predicates and function quantifiers," forthcoming in *Transactions of the American Mathematical Society*.

[13] ———, "Hierarchies of number-theoretic predicates," forthcoming in *Bulletin of the American Mathematical Society*.

[14] S. C. Kleene and Emil L. Post, "The upper semi-lattice of degrees of recursive unsolvability," *Annals of Mathematics*, ser. 2, vol. 59 (1954), pp. 379-407.

[15] W. Markwald, "Zur Theorie der konstruktiven Wohlordnungen," *Mathematische Annalen*, vol. 127 (1954), pp. 135-149.

[16] D. Nelson, "Recursive functions and intuitionistic number theory," *Transactions of the American Mathematical Society*, vol. 61 (1947), pp. 307-368.

[17] R. Péter, "Über den Zusammenhang der verschiedenen Begriffe der rekursiven Funktion," *Mathematische Annalen*, vol. 110 (1934), pp. 612-632.

[18] ———, *Rekursive Funktionen*, Budapest (1951).

[19] E. L. Post, "Recursively enumerable sets of positive integers and their decision problems," *Bulletin of the American Mathematical Society*, vol. 50 (1944), pp. 284-316.

[20] A. M. Turing, "Systems of logic based on ordinals," *Proceedings of the London Mathematical Society*, ser. 2, vol. 45 (1939), pp. 161-228.

LIE GROUPS AND LIE HYPERALGEBRAS OVER A FIELD OF CHARACTERISTIC $p > 0$ (IV).*

By JEAN DIEUDONNÉ.

1. **Introduction.** In this paper, we develop to their full generality the methods introduced in [4] and [5], which enable us to round off, in a way, the theory of abelian Lie groups over a field K. More precisely, we show that this theory is entirely equivalent to the study of finitely generated *modules* of a certain type over a certain ring $\mathfrak{A}$ and of their homomorphisms (Theorems 3 and 5). If, for simplicity's sake, we suppose K perfect, then $\mathfrak{A}$ is defined in the following way: let $\mathfrak{R}$ be the ring of integers in the complete unramified p-adic field $\mathfrak{K}$ having K as its field of residues; then $\mathfrak{A}$ is the ring of *non commutative power series* $\boldsymbol{r}_0 + T\boldsymbol{r}_1 + \cdots + T^n\boldsymbol{r}_n + \cdots$, where $\boldsymbol{r}_n \in \mathfrak{R}$ and $\boldsymbol{r}T = T\boldsymbol{r}^\sigma$ for every $\boldsymbol{r} \in \mathfrak{R}$, σ being the automorphism of $\mathfrak{R}$ which corresponds to the automorphism $\xi \rightarrow \xi^p$ of K in the usual way. The ring $\mathfrak{A}$ is not a principal (noncommutative) ideal ring, and this accounts for the fact that the theory cannot yet be considered as complete, there being, as far as I know, no satisfactory theory of finitely generated modules over such a ring; but at least the kind of difficulties which lie in the way is now clearly delimited.

The main technical tools, as in [5], are the *hyperexponential group* and direct products of finite numbers of such groups; their role as "universal groups" stems from the fundamental fact that their hyperalgebra is *free* (in other words, it is an algebra of polynomials). For technical reasons, it is sometimes easier to work with the (additive) *Witt group,* which is isomorphic to the hyperexponential group, but is more convenient for some computations. It may be hoped that these tools will also one day be useful in the, apparently much more difficult, study of noncommutative formal Lie groups over a field of characteristic p.

2. **The ring of endomorphisms of the Witt group.** Owing to the close relationship between the (additive) Witt group and p-adic rings, it will be technically more convenient to determine first the endomorphisms of the Witt group; the determination of the endomorphisms of any group of hyperexponential type will then be quite trivial, since such a group is isomorphic to the Witt group [5, no. 7].

---

* Received January 20, 1955.

We do not repeat here the definition of the additive Witt group W (see for instance [5, no. 7], or [6, no. 2]); it is a recursive group over the weighted variables x_i $(i \geq 0)$, x_i having the weight p^i. We shall write $\phi = (\phi_0, \phi_1, \cdots, \phi_i, \cdots)$ its group law, $\phi_i(\boldsymbol{x}, \boldsymbol{y})$ being an isobaric polynomial of total weight p^i in $x_0, \cdots, x_i, y_0, \cdots, y_i$; more precisely, one has

$$\phi_i(\boldsymbol{x}, \boldsymbol{y}) = x_i + y_i + \psi_i(x_0, y_0, x_1, y_1, \cdots, x_{i-1}, y_{i-1}),$$

and the coefficients of ψ_i are in the prime field $\boldsymbol{F}_p$.

We can of course consider W as a group over *any* field K of characteristic p; but the notion of endomorphism of W depends on K, and we will say that $\boldsymbol{u} = (u_0, u_1, \cdots, u_i, \cdots)$ is a *K-endomorphism* of W if the coefficients of the formal power series u_i are in K. We will write $\mathcal{E}(K)$, or simply $\mathcal{E}$, for the set of K-endomorphisms of W which are *recursive*, that is, such that $u_i(\boldsymbol{x})$ only contains the indeterminates $x_0, x_1, \cdots, x_i$; on the other hand, we depart here from the conventions made in [5, no. 2] in that *we do not assume* that the minimum weight of u_i tends to $+\infty$ with i; this can be done here, due to the fact that ϕ_i and u_i contain a finite number of indeterminates. The set $\mathcal{E}(K)$ can then be made into a *ring* in the usual fashion, $\boldsymbol{uv}$ meaning the composite endomorphism $\boldsymbol{x} \to \boldsymbol{u}(\boldsymbol{v}(\boldsymbol{x}))$, and $\boldsymbol{u} + \boldsymbol{v}$ the endomorphism $\boldsymbol{w}$ such that $\boldsymbol{w}(\boldsymbol{x})$ is obtained by taking the "product" in W of $\boldsymbol{u}(\boldsymbol{x})$ and $\boldsymbol{v}(\boldsymbol{x})$; in other words, we have

$$w_i(\boldsymbol{x}) = \phi_i(u_0(\boldsymbol{x}), \cdots, u_i(\boldsymbol{x}); v_0(\boldsymbol{x}), \cdots, v_i(\boldsymbol{x})) \qquad (i = 0, 1, \cdots).$$

3. In order to study the structure of the ring $\mathcal{E}(K)$, we first consider certain special types of endomorphisms of W:

1. The "Frobenius homomorphism" $\boldsymbol{p}$, such that

$$(1) \qquad p_i(\boldsymbol{x}) = x_i^p \qquad (i = 0, 1, \cdots).$$

2. The "shift" homomorphism $\boldsymbol{t}$, such that

$$(2) \qquad t_0(\boldsymbol{x}) = 0, \qquad t_i(\boldsymbol{x}) = x_{i-1} \qquad (i = 1, 2, \cdots).$$

It is well-known [7] that the product

$$(3) \qquad \boldsymbol{\pi} = \boldsymbol{tp} = \boldsymbol{pt}$$

is the "p-th-power endomorphism" of the group W, and we have

$$(4) \qquad \pi_0(\boldsymbol{x}) = 0, \qquad \pi_i(\boldsymbol{x}) = x_{i-1}^p \qquad (i = 1, 2, \cdots).$$

3. Let $\mathfrak{W}(K)$ be the ring of *Witt-vectors* $A = (a_0, a_1, \cdots, a_i, \cdots)$ *over*

K; it is a commutative ring of characteristic 0 without zero divisors, with unit $I=(1,0,\cdots,0,\cdots)$ [7]. To each such vector A corresponds an endomorphism

$$x \rightarrow A \cdot x \tag{5}$$

of the additive group W, the "product" being taken in the sense of the Witt *multiplicative* group; $y_i=(A \cdot x)_i$ is an isobaric polynomial in $x_0,\cdots,x_i$, of weight p^i, with coefficients in K. If in y_i we substitute 1 for x_0, 0 for all x_j of index $j \geqq 1$, we obtain the i-th component a_i of A; therefore there is a one-to-one correspondence between the Witt-vectors A and the endomorphisms (5) they define in W, and this correspondence is in fact a ring-isomorphism between $\mathfrak{W}(K)$ and the corresponding subring of $\mathcal{E}(K)$. We shall from now on *identify* these two rings. The endomorphism $\boldsymbol{\pi}$ is then identified with the element $p \cdot I=(0,1,0,\cdots,0,\cdots)$ of $\mathfrak{W}(K)$; $\mathfrak{W}(K)$ is a complete valuation ring, the valuation $\omega(A)$ being the smallest index i such that $a_i \neq 0$; the field of residues is K, and when K is perfect, $\mathfrak{W}(K)$ is isomorphic to the unique unramified complete valuation ring having K as its field of residues [7]. We define a ring-endomorphism σ of $\mathfrak{W}(K)$ by putting

$$A^\sigma=(a_0^p, a_1^p,\cdots,a_i^p,\cdots) \tag{6}$$

With this definition, it is readily verified that we have the commutation laws

$$\boldsymbol{p}A=A^\sigma \boldsymbol{p}, \qquad A\boldsymbol{t}=\boldsymbol{t}A^\sigma. \tag{7}$$

We also notice that the series

$$A_0+A_1\boldsymbol{p}+\cdots+A_k\boldsymbol{p}^k+\cdots \tag{8}$$

$$B_0+\boldsymbol{t}B_1+\boldsymbol{t}^2B_2+\cdots+\boldsymbol{t}^kB_k+\cdots \tag{9}$$

represent meaningful endomorphisms, when the A_k and B_k are arbitrary elements of the ring $\mathfrak{W}(K)$; indeed, the j-th component of $(A_k\boldsymbol{p}^k)\cdot x$ is an isobaric polynomial in $x_0,\cdots,x_j$, of weight p^{j+k}; it follows that, if $\boldsymbol{s}_k$ is the sum of the first k terms of (8), the terms of weight $<p^k$ in any component of $\boldsymbol{s}_k(x)$ are the same as in the component of $\boldsymbol{s}_{k+1}(x)$ of the same index. On the other hand, any component of $(\boldsymbol{t}^kB_k)\cdot x$ of index $\leqq k$ is 0; therefore, if $\boldsymbol{u}_k$ is the sum of the first k terms of (9), the components of $\boldsymbol{u}_k(x)$ of index $\leqq k$ are the same as those of $\boldsymbol{u}_{k+1}(x)$, and this proves our assertion.

It follows from (7) that the series (8) (resp. (9)) is invertible in $\mathcal{E}(K)$ if and only if A_0 (resp. B_0) is invertible in $\mathfrak{W}(K)$, which means that $\omega(A_0)=0$, or equivalently that if $A_0=(a_{00},a_{01},\cdots,a_{0i},\cdots)$, we have $a_{00}\neq 0$ in K. The inverse is computed in the usual way, as another power series of type (8), whose coefficients are determined recursively by those of (8). The results are similar for (9).

4. We can now characterize the elements of $\mathcal{E}(K)$:

THEOREM 1. *If K is perfect, every endomorphism $\boldsymbol{u} \varepsilon \mathcal{E}(K)$ can be written in one and only one way in the form*

$$\boldsymbol{u} = \sum_{k=0}^{\infty} A_k \boldsymbol{p}^k + \sum_{k=1}^{\infty} \boldsymbol{t}^k B_k \tag{10}$$

where the A_k and B_k are elements of the ring $\mathfrak{W}(K)$.

Writing that $\boldsymbol{u}$ is an endomorphism, we obtain first the relation

$$u_0(x_0 + y_0) = u_0(x_0) + u_0(y_0)$$

from which the form of u_0 follows trivially:

$$u_0(x_0) = a_0 x_0 + a_1 x_0^p + \cdots + a_k x_0^{p^k} + \cdots.$$

Let $A_i = (a_i, 0, 0, \cdots)$ for each i, and consider the endomorphism $\boldsymbol{v}_0 = A_0 + A_1 \boldsymbol{p} + \cdots + A_k \boldsymbol{p}^k + \cdots$; it is clear that $u_0(x_0) = v_{00}(x_0)$, and therefore the endomorphism $\boldsymbol{w} = \boldsymbol{u} - \boldsymbol{v}_0$ is such that $w_0(x_0) = 0$.

As a basis for an inductive argument, let us now suppose that we have $u_j(\boldsymbol{x}) = 0$ for $j < i$, and let us determine $u_i(x)$. We must have

$$\begin{aligned} u_i(x_0, \cdots, x_i) &+ u_i(y_0, \cdots, y_i) \\ &= u_i(x_0 + y_0, \phi_1(\boldsymbol{x}, \boldsymbol{y}), \cdots, \phi_i(\boldsymbol{x}, \boldsymbol{y})). \end{aligned} \tag{11}$$

On the other hand, we have $\boldsymbol{\pi u} = \boldsymbol{u \pi}$, and this, using the inductive hypothesis and relations (4), gives

$$u_i(0, x_0^p, \cdots, x_{i-1}^p) = 0, \tag{12}$$

in other words

$$u_i(0, x_1, \cdots, x_i) = 0. \tag{13}$$

Therefore every monomial in u_i contains x_0; let us show that u_i *does not contain* $x_1, \cdots, x_i$. Indeed, suppose $u_i(\boldsymbol{x}) = \sum_{r=0}^{\infty} x_i^r f_r(x_0, \cdots, x_{i-1})$, and suppose h is the smallest value of $r > 0$ for which $f_r \neq 0$; identifying in both sides of (11) the terms in x_i^h gives

$$f_h(x_0, \cdots, x_{i-1}) = f_h(x_0 + y_0, \cdots, \phi_{i-1}(\boldsymbol{x}, \boldsymbol{y})) + F$$

where F is a polynomial in $x_0, y_0, \cdots, x_{i-1}, y_{i-1}, y_i$, each term of which contains either y_i or an x_k of index $k \leq i-1$ (due to the fact that every monomial in ψ_i contains at least such an x_k). If we replace $x_0, \cdots, x_{i-1}$ by 0 in this identity, we get therefore

$$c_h = f_h(y_0, \cdots, y_{i-1}) + \sum_{k=1}^{\infty} y_i^k g_k(y_0, \cdots, y_{i-1})$$

where c_h is the constant term in f_h, and this shows that f_h is a *constant*. But as we have seen before that each monomial in f_h must contain x_0, we have $f_h = 0$, hence u_j cannot contain x_i. Suppose then that k is the highest index j such that u_i contains x_j, and let again $u_i(\boldsymbol{x}) = \sum_{r=0}^{\infty} x_k{}^r g_r(x_0, \cdots, x_{k-1})$; suppose $k > 0$, and that h is the smallest value of $r > 0$ for which $g_r \neq 0$. Identifying on both sides of (11) the terms in $x_k{}^h$, we see by the same argument as above that $g_h = 0$. Equation (11) reduces therefore to $u_i(x_0 + y_0) = u_i(x_0) + u_i(y_0)$, hence

$$u_i(\boldsymbol{x}) = c_0 x_0 + c_1 x_0{}^p + \cdots + c_k x_0{}^{p^k} + \cdots.$$

Let $C_i = (c_i, 0, 0, \cdots)$ in $\mathfrak{W}(K)$, and consider the endomorphism

$$\boldsymbol{v}_i = \boldsymbol{t}^i (C_0 + C_1 \boldsymbol{p} + \cdots + C_j \boldsymbol{p}^j + \cdots). \tag{14}$$

It is clear that the j-th components of $\boldsymbol{u}(\boldsymbol{x})$ and $\boldsymbol{v}_i(\boldsymbol{x})$ are the same for $j \leq i$ (and 0 for $j < i$). The inductive argument can then be carried on in an obvious fashion, and shows that

$$\boldsymbol{u} = \boldsymbol{v}_0 + \boldsymbol{v}_1 + \cdots + \boldsymbol{v}_i + \cdots$$

the infinite sum having a meaning in $\mathcal{E}(K)$. Therefore, we see that we can write $\boldsymbol{u} = \sum_{i,j} \boldsymbol{t}^i C_{ij} \boldsymbol{p}^j$, where $C_{ij} \varepsilon \mathfrak{W}(K)$. If we now assume that K is *perfect*, we have

$$\boldsymbol{t}^i C_{ij} \boldsymbol{p}^j = \boldsymbol{t}^{i-j} C_{ij}{}^{\sigma^{-j}} \boldsymbol{\pi}^j \qquad \text{if } j \leq i$$

$$\boldsymbol{t}^i C_{ij} \boldsymbol{p}^j = C_{ij}{}^{\sigma^{-i}} \boldsymbol{\pi}^i \boldsymbol{p}^{j-i} \qquad \text{if } j \geq i,$$

from which (using the fact that $\mathfrak{W}(K)$ is complete) it is easy to deduce that $\boldsymbol{u}$ has the form (10).

To prove uniqueness, suppose $\boldsymbol{u}$ is equal to the right-hand side of (10) and is 0; we have to show that all the A_k and B_k are 0. Suppose the contrary, and let us prove first that it is impossible that $\omega(A_k) = 0$ or $\omega(B_k) = 0$ for some k. Indeed, suppose first $\omega(A_k) = 0$ for some $A_k \neq 0$, and let h be the smallest of the indices k having that property; then if $A_h = (a_{h0}, \cdots)$, $a_{h0} \neq 0$ and the component of index 0 of $\boldsymbol{u}(\boldsymbol{x})$ would be $a_{h0} x_0{}^{p^h} + \cdots$, the terms omitted having degree $> p^h$, contrary to assumption. If on the other hand

$\omega(B_k)=0$ for some $B_k \neq 0$, let again $h \geqq 1$ be the smallest index having that property; then, in the component of $\boldsymbol{u}(\boldsymbol{x})$ of index h, there would be only one term $b_{h0}x_0$ of weight 0, with $b_{h0} \neq 0$, which is again contrary to the assumption $\boldsymbol{u}=0$. Suppose now we have proved that the minimum value of the $\omega(A_k)$ and $\omega(B_k)$ (for the non zero coefficients) cannot be $< j$, and let us prove it cannot be equal to j. Let us write $A_k = A_k'\boldsymbol{\pi}$, $B_k = B_k'\boldsymbol{\pi}$, where the minimum value of the $\omega(A_k')$ and $\omega(B_k')$ is therefore $j-1$. As $\boldsymbol{\pi} = \boldsymbol{tp}$, and the relation $\boldsymbol{wp}=0$ implies $\boldsymbol{w}=0$, we have

$$\sum_{k=1}^{\infty} \boldsymbol{t}^{k+1} B_k'^{\sigma} + \boldsymbol{t} A_0'^{\sigma} + A_1'\boldsymbol{\pi} + \sum_{k=2}^{\infty} A_k'\boldsymbol{\pi}\boldsymbol{p}^{k-1} = 0.$$

This already implies contradiction if $\omega(A_0')$ or one of the $\omega(B_k')$ is $j-1$; if not, we can again divide by $\boldsymbol{p}$ on the right and after a finite number of steps, we reach the desired contradiction.

Theorem 1 is thus completely proved. In what follows, we shall use only the endomorphisms of type

$$B_0 + \boldsymbol{t}B_1 + \cdots + \boldsymbol{t}^k B_k + \cdots \tag{15}$$

which (without any assumption on K) constitute a subring $\mathcal{E}^+(K)$ of $\mathcal{E}(K)$. The preceding argument shows that an endomorphism of $\mathcal{E}^+(K)$ has only one development (15).

5. Abelian groups associated with endomorphisms of H^n. There exists a recursive isomorphism $\boldsymbol{h}$ of the additive Witt group W on the hyperexponential group H, such that $h_i(\boldsymbol{x})$ is an isobaric polynomial $x_i + \cdots$ of weight p^i for each $i \geqq 0$, $h_0(\boldsymbol{x}) = x_0$, $h_1(\boldsymbol{x}) = x_1$ ([5, no. 7], and [6]). Any K-endomorphism of H (in the same sense as in no. 2) can therefore be written $\boldsymbol{huh}^{-1}$, where $\boldsymbol{u} \in \mathcal{E}(K)$. In what follows, we will identify $\boldsymbol{u}$ and $\boldsymbol{huh}^{-1}$; this will certainly lead to no confusion regarding the endomorphisms $\boldsymbol{t}$ and $\boldsymbol{p}$, for $\boldsymbol{hph}^{-1}$ and $\boldsymbol{hth}^{-1}$ are respectively the Frobenius homomorphism and the shift homomorphism in the group H.

Consider now the endomorphisms of the group H^n, direct product of n groups isomorphic to H; we will denote by $\boldsymbol{x}^{(i)} = (x_0^{(i)}, \cdots, x_h^{(i)}, \cdots)$ $(1 \leqq i \leqq n)$ the systems of indeterminates corresponding to each of the factors. A recursive endomorphism $\boldsymbol{u}$ of H^n will therefore consist in n systems $\boldsymbol{u}^{(i)} = (u_0^{(i)}, \cdots, u_h^{(i)}, \cdots)$ of power series in the $x_h^{(i)}$ $(1 \leqq i \leqq n)$, such that $u_h^{(i)}$ contains only the $x_k^{(j)}$ such that $k \leqq h$ and $1 \leqq j \leqq n)$. The standard argument determining the endomorphisms of a module which is a direct sum of n isomorphic submodules proves in addition that the endomorphisms $\boldsymbol{u}$ correspond, in a one-to-one fashion, to *square matrices* $(f^{(ij)})$ of order n, where the $f^{(ij)}$ $(1\leq i\leq n, 1\leq j\leq n)$ are elements of the ring of endomorphisms of H (identified to $\mathcal{E}(K)$). More precisely, one can write symbolically

$$(16)\qquad \boldsymbol{u}^{(i)}(\boldsymbol{x}^{(1)},\cdots,\boldsymbol{x}^{(n)})=\sum_{j=1}^{n}f^{(ij)}(\boldsymbol{x}^{(j)})$$

the "sum" in the right hand side being understood as taken *for the group law of H.*

From now on until no. 10, $\phi=(\phi_h)_{h=0,1\cdots}$ will always designate the group law of H, so that, with the notations of [5, no. 5], we have

$$\phi_h(x_0,\cdots,x_h;y_0,\cdots,y_h)=x_h+y_h+\sum_{1<k<p^h}E_k(\boldsymbol{x})E_{p^h-k}(\boldsymbol{y})$$

where E_k is an isobaric polynomial of weight k.

6. With the preceding notations, we now suppose given an endomorphism u of H^n such that the endomorphisms $f^{(ij)}$ of H *belong to the ring* $\mathcal{E}^{+}(K)$ (no. 4); owing to (16), this implies that in the power series $u_h^{(i)}$, all monomials *have a weight* $\leq p^h$, and in particular the terms containing the indeterminates $x_h^{(j)}$ $(1\leq j\leq n)$ are of the *first* (*total*) *degree.* Generalizing the method introduced in [5, no. 9], we are going to define, with the help of $\boldsymbol{u}$, the *hyperalgebra of an abelian group of dimension* n (over K).

In order to do this, we consider the algebra P of polynomials with respect to n sequences of indeterminates S_{hi} $(1\leq i\leq n, h=0,1,\cdots)$, having coefficients in K, and in this algebra, we consider the ideal $\mathfrak{a}$ generated by the polynomials

$$(17)\qquad S_{hi}^{p}-u_h^{(i)}(S_{01},\cdots,S_{0n};\cdots;S_{h1},\cdots,S_{hn})$$
$$(1\leq i\leq n;h=0,1,\cdots).$$

For each index $\alpha=(\alpha_1,\cdots,\alpha_n)$ (α_i integers ≥ 0), let

$$E_{\alpha}(S_{01},\cdots,S_{h1},\cdots;S_{02},\cdots,S_{h2},\cdots;\cdots;S_{0n},\cdots,S_{hn},\cdots)$$
$$=E_{\alpha_1}(S_{01},\cdots,S_{h1},\cdots)E_{\alpha_2}(S_{02},\cdots,S_{h2},\cdots)\cdots E_{\alpha_n}(S_{0n},\cdots,S_{hn},\cdots).$$

Let X_{hi} be the class of S_{hi} in the quotient algebra $P/\mathfrak{a}$, and

$$(18)\qquad Z_{\alpha}=E_{\alpha}(X_{01},\cdots,X_{h1},\cdots;\cdots;X_{0n},\cdots,X_{hn},\cdots).$$

THEOREM 2. *The elements Z_α form a basis of the algebra $P/\mathfrak{a}$, such that, with this basis, $P/\mathfrak{a}$ is the hyperalgebra of an abelian group G of dimension n.*

It is clear that the algebra $P/\mathfrak{a}$ is generated by the X_{hi}, and more precisely, from the relations

$$(19) \qquad X_{hi}^p = u_h^{(i)}(X_{01}, \cdots, X_{0n}; \cdots; X_{h1}, \cdots, X_{hn})$$

it follows, by an easy inductive argument on h, that the elements of $P/\mathfrak{a}$ are linear combinations of monomials X_α in the X_{hi} where the exponent of each X_{hi} is $< p$; moreover, relations (19) also show that the linear combinations of the X_α of height $h(\alpha) < r$ constitute a subalgebra $\mathfrak{S}_r$ of $P/\mathfrak{a}$. From the definition (18), it follows that the Z_α belong to $\mathfrak{S}_r$ for $h(\alpha) < r$; we are going to prove that the Z_α of height $h(\alpha) < r$ constitute a *basis* of $\mathfrak{S}_r$; the family (Z_α) (without restriction on α) will then of course constitute a basis of $P/\mathfrak{a}$.

The first step consists in proving that the

$$X_\alpha = X_{01}^{\nu_{01}} \cdots X_{0n}^{\nu_{0n}} X_{11}^{\nu_{11}} \cdots X_{1n}^{\nu_{1n}} \cdots X_{h1}^{\nu_{h1}} \cdots X_{hn}^{\nu_{hn}}$$

(where $\alpha_i = \sum_{k=0}^{h} \nu_{ki} p^k$ is the p-adic development of α_i, and $h < r$) constitute a basis for $\mathfrak{S}_r$. This will be proved by induction on r, as a consequence of the fact that all the monomials, in the polynomial $u_h^{(i)}$, have a weight $\leq p^h$, and of the elementary lemma:

Lemma 1. *Let A be a commutative ring with unit element, and let $\mathfrak{b}$ be the ideal in the ring of polynomials $A[Y_1, \cdots, Y_n]$, generated by the polynomials*

$$Q_i = Y_i^m - \sum_{j=1}^{n} b_{ij} Y_j - c_i. \qquad (1 \leq i \leq n)$$

m being an integer > 1, the b_{ij} and c_i elements of A. Then if $\bar{Y}_i$ is the class of $Y_i \bmod \mathfrak{b}$, the monomials $\bar{Y}_1^{\mu_1} \bar{Y}_2^{\mu_2} \cdots \bar{Y}_n^{\mu_n}$ with $0 \leq \mu_i < m$, constitute a basis of the quotient ring $A[Y_1, \cdots, Y_n]/\mathfrak{b}$ over the ring A.

To prove the lemma, let us assign weight 1 to the Y_i, weight m to each polynomial Q_i. Then it is readily seen by induction on k that every monomial in $Y_1, \cdots, Y_n$, of total degree $\leq k$, can be expressed as a linear combination of monomials in $Y_1, \cdots, Y_n, Q_1, \cdots, Q_n$, of total weight $\leq k$, the exponent of each Y_i in these monomials being $< m$. As the number of these last monomials of total weight $\leq k$ is exactly equal to the number of monomials in $Y_1, \cdots, Y_n$ of total degree $\leq k$, they constitute a basis for the set of polynomials in $Y_1, \cdots, Y_n$ of total degree $\leq k$; the lemma follows immediately from this result.

The fact that the X_α with $h(\alpha) < r$ constitute a basis for $\mathfrak{S}_r$ is then proved by applying the lemma in succession to $A = K$, $A = \mathfrak{S}_1, \cdots, A = \mathfrak{S}_{r-1}$.

7. Next we remark (see [5, no. 8]) that if $k = \nu_0 + \nu_1 p + \cdots + \nu_h p^h$ (p-adic development, $\nu_h > 0$), there is in $E_k(x_0, \cdots, x_h)$ the monomial $\dfrac{x_0^{\nu_0} x_1^{\nu_1} \cdots x_h^{\nu_h}}{\nu_0! \nu_1! \cdots \nu_h!}$; moreover, if we order the monomials in $x_0, \cdots, x_h$ in the *inverse* lexicographic order, the preceding monomial is the "first" in E_k. Let us next order the S_{hi} in a linear sequence:

$$S_{01}, \cdots, S_{0n}, S_{11}, \cdots, S_{1n}, \cdots, S_{h1}, \cdots, S_{hn}, \cdots.$$

If we put $\alpha_i = \sum_l \nu_{li} p^l$ (p-adic development), the preceding remark shows that if we order the monomials in the S_{hi} in the inverse lexicographic order, the "first" monomial in E_α is the monomial $\dfrac{1}{\alpha!} S_\alpha$ (with notation similar to that of the X_α). Consider now the expression of Z_α as a linear combination of the monomials X_λ, and let us order these monomials in the same order as the corresponding monomials in the S_{hi}. We observe that if in a monomial $S_{01}^{\mu_{01}} \cdots S_{hn}^{\mu_{hn}}$ in E_α there is an exponent $\mu_{kj} \geq p$, then there must be an index $l > k$ such that we will have $\mu_{lj} < \nu_{lj}$, owing to the fact that the E_k are *isobaric*; considering the *largest* index k for which there is a j such that $\mu_{kj} \geq p$, we see therefore that in the expression of the corresponding monomial $X_{01}^{\mu_{01}} \cdots X_{hn}^{\mu_{hn}}$ as a linear combination of the X_λ, every monomial X_λ is *after* the monomial $\dfrac{1}{\alpha!} X_\alpha$ (which itself comes from the replacement of the S_{hi} by the X_{hi} in $\dfrac{1}{\alpha!} S_\alpha$), due to relations (19). In other words, in the expression of Z_α as a linear combination of the X_λ, the "first" term is $\dfrac{1}{\alpha!} X_\alpha$. From the linear independence of the X_α, it follows at once that the Z_α are linearly independent over K.

This result ends our proof of the fact that the Z_α with $h(\alpha) < r$ constitute a basis for $\mathfrak{Z}_r$, for their number is the same as that of the X_α with $h(\alpha) < r$.

8. We can now write

$$Z_\alpha Z_\beta = \sum_\gamma d_{\alpha\beta\gamma} Z_\gamma \qquad (d_{\alpha\beta\gamma} \in K) \tag{20}$$

and by definition, these relations can be interpreted as polynomial congruences

$$E_\alpha E_\beta \equiv \sum_\gamma d_{\alpha\beta\gamma} E_\gamma \qquad (\text{mod. } \mathfrak{a}) \tag{21}$$

in the ring P. To prove Theorem 2, we apply the existence theorem [5, Prop. 1]; replace in P and in the polynomials (17) the indeterminates S_{hi}

by new indeterminates T_{hi}, obtaining thus a ring P_1 of polynomials and an ideal $\mathfrak{a}_1$ in P_1, and let Y_{hi} be the class of $T_{hi}\,\mathrm{mod.}\,\mathfrak{a}_1$. We then form the elements

$$Z_\alpha{}^0 = \sum_{0\leq\beta\leq\alpha} E_\beta(X_{01},\cdots,X_{hn},\cdots)E_{\alpha-\beta}(Y_{01},\cdots,Y_{hn},\cdots)$$

and we have to prove that these elements satisfy the *same* relations (20) as the Z_α. But from the expression of the ϕ_h in terms of the E_k and the definition of the hyperexponential series, it follows that we have [5, no. 8]

$$Z_\alpha{}^0 = E_\alpha(\Phi_{01},\cdots,\Phi_{h1},\cdots;\cdots;\Phi_{0n},\cdots,\Phi_{hn},\cdots) \tag{22}$$

where

$$\Phi_{hi} = \phi_h(X_{0i}, Y_{0i}, X_{1i}, Y_{1i},\cdots,X_{hi}, Y_{hi}).$$

Now, if in the polynomial congruences (21) we replace each S_{hi} by Φ_{hi}, we see that the difference

$$Z_\alpha{}^0 Z_\beta{}^0 - \sum_\gamma d_{\alpha\beta\gamma} Z_\gamma{}^0$$

is a linear combination (with coefficients which are polynomials in the Φ_{hi}, and therefore in the X_{hi} and Y_{hi}) of the polynomials

$$\Phi_{hi}{}^p - u_h{}^{(i)}(\Phi_{01},\cdots,\Phi_{0n};\cdots;\Phi_{h1},\cdots,\Phi_{hn}) \tag{23}$$

and therefore our proof will be ended if we establish that the expressions (23) all vanish. But as the ϕ_h have their coefficients in the prime field $\boldsymbol{F}_p$, we have

$$\Phi_{hi}{}^p = \phi_h(X_{0i}{}^p, Y_{0i}{}^p,\cdots,X_{hi}{}^p, Y_{hi}{}^p)$$

and if we equal the expression (23) to 0, we obtain the relation

$$\begin{aligned}&\phi_h(X_{0i}{}^p, Y_{0i}{}^p,\cdots,X_{hi}{}^p, Y_{hi}{}^p)\\ &\quad = u_h{}^{(i)}(\phi_0(X_{01}, Y_{01}),\cdots,\phi_0(X_{0n}, Y_{0n});\cdots;\\ &\qquad\quad \cdots;\phi_h(X_{01},\cdots,Y_{h1}),\cdots,\phi_h(X_{0n},\cdots,Y_{hn})).\end{aligned}$$

But if we replace the $X_{hi}{}^p$ by their expressions (19), and the $Y_{hi}{}^p$ by the corresponding expressions, the relations we obtain follow from the assumption that $\boldsymbol{u}$ is an *endomorphism* of the group H^n. The proof of Theorem 2 is thus complete.

9. We will say that the abelian group G, the existence of which follows from Theorem 2, is *associated* to the endomorphism $\boldsymbol{u}$ of H^n. Our proof gives an "explicit" determination of that group when $\boldsymbol{u}$ is known; more precisely, the translation operator in G is given by

$$R_y = \prod_{i=1}^{n} Hex(y_i X_{0i}, y_i{}^p X_{1i},\cdots,y_i{}^{p^h} X_{hi}\cdot\cdots)$$

where in the development of the right-hand side, the powers of the X'_{hi} must be reduced to exponents $< p$ according to the relations (19). Moreover, the algebra P can be identified with the hyperalgebra of the product U^n of n groups isomorphic with the "universal group" U [5, no. 8], the S_{hi} being the invariant semi-derivations which generate that hyperalgebra; the proof of Theorem 2 shows then [5, Prop. 2] that the natural mapping of P onto $P/\mathfrak{a}$ is the derived homomorphism $\boldsymbol{v}'$ of a homomorphism $\boldsymbol{v}$ of U^n onto G.

10. Some examples. It is easy to reformulate in terms of matrices (with elements in $\mathcal{E}^+(K)$) the definitions of the particular groups considered in [4] and [5]. For instance, to the null matrix (of order n) corresponds the direct product $W_1{}^n$ of n additive groups; to the unit matrix. the direct product $(W_1{}^*)^n$ of n multiplicative groups. The group $G_{n,m,r}$ defined in [5. no. 9] corresponds to the matrix

$$\begin{pmatrix} 0 & I & 0 & & \cdots & 0 \\ 0 & 0 & I & & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & & \cdots & I \\ 0 & \cdots & & \boldsymbol{t}^r & \cdots & 0 \end{pmatrix}$$

where in the last row, the element $\boldsymbol{t}^r$ is in the $(m+1)$-st column.

Similarly, the "streak", "tree" and "braid" types of groups of dimension 2 defined in [5, no. 10], correspond respectively to the matrices

$$\begin{pmatrix} 0 & \boldsymbol{t}^r \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & \boldsymbol{t}^r \\ 0 & \boldsymbol{t}^s \end{pmatrix}, \qquad \begin{pmatrix} 0 & \boldsymbol{t}^r \\ \boldsymbol{t}^s & 0 \end{pmatrix}.$$

11. The structure of abelian formal Lie groups. The construction of abelian Lie groups given by Theorem 2 is in fact general enough to give *all* abelian Lie groups, up to an isomorphism; in other words:

THEOREM 3. *Every abelian formal Lie group G of dimension n over a field K is isomorphic to an abelian group associated to an endomorphism of H^n.*

We start by remarking that the invariant derivations X_{0i} $(1 \leq i \leq n)$ of G satisfy relations of the form

$$X_{0i}{}^p = \sum_{j=1}^{n} a_{ij} X_{0j} \qquad (1 \leq i \leq n,\ a_{ij} \in K).$$

We consider the Witt vectors $A_{ij} = (a_{ij}, 0, 0, \cdots)$, and the endomorphism $\boldsymbol{u}^{(0)}$ of H^n defined by the matrix (A_{ij}) (whose elements are identified with endomorphisms of H, as explained in no. 5). Let $L^{(0)}$ be the abelian group asso-

and

$$a_{ij}^{p^{r+1}} = \sum_k g_{i,k,l\epsilon_j} g_{i,p^{r+1}-k,(p^{r+1}-l)\epsilon_j}$$

for any ρ such that $0 < \rho < \mu$ and any l such that $0 < l < p^{r+1}$; in other words, the $b'_{i\mu}$ for $|\mu| > 1$ are entirely determined by the $g_{ik\rho}$. But from our assumptions it follows that we also have

$$(28) \qquad Z_{k\epsilon_i}{}^p = \sum_\rho g_{ik\rho} Z_\rho \qquad (1 \leq i \leq n, 0 < k < p^{r+1})$$

with the *same* coefficients. On the other hand, we have by [4, Lemma 2]

$$(29) \quad X_{r+1,i}{}^p(fg) = f \cdot X_{r+1,i}{}^p g + g \cdot X_{r+1,i}{}^p f \\ + \sum_{0<k<p^{r+1}} (Z_{k\epsilon_i}{}^p f)(Z_{(p^{r+1}-k)\epsilon_i}{}^p g).$$

Replacing the $Z_{k\epsilon_i}{}^p$ by their expressions (28), and using relations (27), we see that in the expression of $Y_i(fg)$ as a sum $\sum_{\lambda,\mu} h_{\lambda\mu}(Z_\lambda f)(Z_\mu g)$, there is no term in which *both* λ and μ are $\neq 0$ and this proves our assertion.

13. We can therefore write

$$Y_i = \sum_{j=1}^n a_{ij}^{(r+1)} X_{0j} \qquad (1 \leq i \leq n, a_{ij}^{(r+1)} \varepsilon K).$$

Suppose the endomorphism $\boldsymbol{u}^{(r)}$ of H^n corresponded to the matrix $(\boldsymbol{f}^{(i,j,r)})$, where $\boldsymbol{f}^{(i,j,r)}$ belongs to $\mathcal{E}^+(K)$. Consider the Witt vectors $A_{ij}^{(r+1)} = (a_{ij}^{(r+1)}, 0, 0, \cdots)$, and the matrix

$$(30) \qquad (\boldsymbol{f}^{(i,j,r+1)}) = (\boldsymbol{f}^{(i,j,r)} + \boldsymbol{t}^{r+1} A_{ij}^{(r+1)})_{1 \leq i,j \leq n}.$$

This matrix corresponds to an endomorphism $\boldsymbol{u}^{(r+1)}$ of H^n, to which in turn correspond an abelian group $L^{(r+1)}$; if we recall the definition of $L^{(r+1)}$ by means of relations (19), we see that we have finally reached the following conclusion: the image of G under the isomorphism $\boldsymbol{q}^{(r)}$, and the group $L^{(r+1)}$, satisfy conditions 1., 2., 3. *where r has been replaced by $r+1$.*

It is now easy to bring our inductive argument to a close. From the last remark of no. 11, it follows that the "infinite product $\cdots \boldsymbol{q}^{(r)} \cdots \boldsymbol{q}^{(1)} \boldsymbol{q}^{(0)}$ is meaningful and represents an isomorphism of G on a group $\bar{G}$. On the other hand, formulae (30) show that the matrices $(\boldsymbol{f}^{(i,j,r)})$ "converge" to a matrix $(\boldsymbol{f}^{(ij)})$ with elements in $\mathcal{E}^+(K)$, corresponding to an endomorphism $\boldsymbol{u}$ of H^n. If $\bar{L}$ is then the abelian group defined by that endomorphism, our construction is such that the constants $d_{\alpha\beta\gamma}$ in the hyperalgebras of $\bar{G}$ and $\bar{L}$ are the *same,* and therefore these two groups are *identical,* which concludes the proof of Theorem 3.

We add two remarks: the first is that, forming (according to Theorem 2) abelian groups corresponding to endomorphisms of H^n, we can always limit ourselves to the case in which the matrix $(\boldsymbol{f}^{(ij)})$ is such that, if $\boldsymbol{f}^{(ij)}=\sum_{k=0}^{\infty}\boldsymbol{t}^k A_{ij}^{(k)}$, the Witt vectors $A_{ij}^{(k)}$ which are $\neq 0$ are such that $\omega(A_{ij}^{(k)})=0$; this follows from (30).

On the other hand, Theorems 2 and 3 yield, for abelian groups, the following result, which is analogous to the well-known theorem in classical Lie theory, asserting that one-parameter subgroups "fill up" the group locally: for any element $X=\sum_{i=1}^{n}\xi_i X_{0i}\neq 0$ $(\xi_i\,\varepsilon\, K)$ of the Lie algebra $\mathfrak{g}_0$ of an abelian group G, there exists a homomorphism $\boldsymbol{u}$ of the hyperexponential group H into G such that $\boldsymbol{u}'(S_{00})=X$ (S_{hi} being the usual generators of the Lie hyperalgebra of H, noted X_{hi} in [5, no. 5]). One has merely to make a linear transformation on the indeterminates in G, which transforms X into X_{00}, for instance; we have then a homomorphism of U^n onto G (no. 9), and the restriction of that homomorphism to the first factor of the product, composed with an isomorphism of H onto U, satisfies the required conditions. It would be very interesting to know if this result extends to non-abelian formal Lie groups.

14. Homomorphisms of abelian groups. It follows from theorem 3 that the study of abelian formal Lie groups can be reduced to that of the groups associated with endomorphisms of groups H^n; we are going to study the homomorphisms of such groups.

Let G and $\bar{G}$ be two such groups, of respective dimensions n and m, and suppose they are respectively associated with an endomorphism $\boldsymbol{u}$ of H^n and an endomorphism $\boldsymbol{v}$ of H^m; $\boldsymbol{u}$ and $\boldsymbol{v}$ can respectively be represented by square matrices $\boldsymbol{U}=(\boldsymbol{f}^{(ij)})$ and $\boldsymbol{V}=(\boldsymbol{g}^{(ij)})$ of order n and m respectively, having their elements in the ring $\mathcal{E}^+(K)$. We shall designate by P and Q the algebras of polynomials, by $\mathfrak{a}$ and $\mathfrak{b}$ the ideals in these algebras determined respectively by $\boldsymbol{u}$ and $\boldsymbol{v}$, so that the hyperalgebras of G and $\bar{G}$ are respectively isomorphic to $P/\mathfrak{a}$ and $Q/\mathfrak{b}$ (no. 6); S_{hi} and T_{hi} will be the indeterminates in P and Q respectively, X_{hi} and Y_{hi} their classes mod. $\mathfrak{a}$ and mod. $\mathfrak{b}$ respectively.

We will first consider the following question. Let $\boldsymbol{w}=(\boldsymbol{w}^{(i)})_{1\leq i\leq n}$ be a homomorphism of H^m into H^n, $\boldsymbol{w}^{(i)}=(w_0^{(i)},\cdots,w_h^{(i)},\cdots)$ being n systems of power series in the $x_h^{(j)}$ $(1\leq j\leq m)$ such that $w_h^{(i)}$ contains only the $x_k^{(j)}$ such that $k\leq h$. To this homomorphism is again associated an $n\times m$

matrix $\boldsymbol{W} = (\boldsymbol{h}^{(ij)})$ $(1 \leq i \leq n, 1 \leq j \leq m)$ with elements $\boldsymbol{h}^{(ij)}$ in $\mathcal{E}(K)$, such that (with the same conventions as in no. 5)

$$\boldsymbol{w}^{(i)}(\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(m)}) = \sum_{j=1}^{m} \boldsymbol{h}^{(ij)}(\boldsymbol{x}^{(j)}) \qquad (1 \leq i \leq n).$$

We will always suppose, in addition, that the $\boldsymbol{h}^{(ij)}$ belong to $\mathcal{E}^{+}(K)$ (which insures that the $w_h^{(i)}$ are in fact polynomials).

We now ask under which conditions on $\boldsymbol{W}$ the formulae

$$s'(X_{hi}) = w_h^{(i)}(Y_{01}, \cdots, Y_{0m}; \cdots; Y_{h1}, \cdots, Y_{hm}) \tag{31}$$
$$(1 \leq i \leq n; h = 0, 1, \cdots)$$

define the derivative $\boldsymbol{s}'$ of a *homomorphism* of G into $\bar{G}$.

THEOREM 4. *Suppose K is a perfect field. In order that formulae* (31) *define a homomorphism $\boldsymbol{s}$ of G into $\bar{G}$, it is necessary and sufficient that there exist an $n \times m$ matrix $\boldsymbol{L}$ with elements in $\mathcal{E}^{+}(K)$, such that*

$$\boldsymbol{UW} - \boldsymbol{W}^{\sigma}\boldsymbol{V} = \boldsymbol{L}(\boldsymbol{\pi} \cdot \boldsymbol{I} - \boldsymbol{t} \cdot \boldsymbol{V}) \tag{32}$$

($\boldsymbol{I}$ being the $m \times m$ unit matrix, and $\boldsymbol{W}^{\sigma}$ being the matrix $((\boldsymbol{h}^{(ij)})^{\sigma})$ the isomorphism σ being defined by $\boldsymbol{a}^{\sigma} = A_0^{\sigma} + \boldsymbol{t}A_1^{\sigma} + \cdots + \boldsymbol{t}^k A_k^{\sigma} + \cdots$ if $\boldsymbol{a} = A_0 + \boldsymbol{t}A_1 + \cdots + \boldsymbol{t}^k A_k + \cdots$).

15. The system of polynomials

$$w_h^{(i)}(T_{01}, \cdots, T_{0m}; \cdots; T_{h1}, \cdots, T_{hm})$$

can be represented by the "one-column matrix" $\boldsymbol{W} \cdot \boldsymbol{x}$, the indeterminates $x_h^{(i)}$ being replaced by T_{hi}. If there is a homomorphism $\boldsymbol{s}$ verifying (31), we must have

$$s'(X_{hi}^{p}) = (w_h^{(i)})^{\sigma}(Y_{01}^{p}, \cdots, Y_{0m}^{p}; \cdots; Y_{h1}^{p}, \cdots, Y_{hm}^{p}) \tag{33}$$

$(w_h^{(i)})^{\sigma}$ being the polynomial obtained by raising all the coefficients of $w_h^{(i)}$ to the power p. But considering the definition of $\bar{G}$ by means of the matrix $\boldsymbol{V}$ (no. 6), we see that the right hand sides of (33) are obtained by replacing T_{hj} by Y_{hj} in the system of polynomials represented by the one-column matrix $(\boldsymbol{W}^{\sigma}\boldsymbol{V}) \cdot \boldsymbol{x}$. On the other hand, the definition of G by means of $\boldsymbol{U}$ and formulae (31) show that (if $\boldsymbol{s}$ exists) the left hand sides of (33) are obtained by replacing T_{hj} by Y_{hj} in the system of polynomials represented by the one-column matrix $(\boldsymbol{UW}) \cdot \boldsymbol{x}$. Therefore, a necessary condition is that the matrix $\boldsymbol{C} = \boldsymbol{UW} - \boldsymbol{W}^{\sigma}\boldsymbol{V}$ be such that, when each T_{hj} is replaced by Y_{hj} in the system of polynomials represented by the one-column matrix $\boldsymbol{C} \cdot \boldsymbol{x}$, the expressions obtained all vanish. This can also be expressed in the following way: the polynomials of the system represented by $\boldsymbol{C}\cdot\boldsymbol{x}$ must *belong to the ideal* $\mathfrak{b}$.

This is certainly the case if $\boldsymbol{C}=\boldsymbol{L}(\boldsymbol{\pi}\cdot\boldsymbol{I}-\boldsymbol{t}\cdot\boldsymbol{V})$, for $\boldsymbol{\pi}\cdot\boldsymbol{I}-\boldsymbol{t}\cdot\boldsymbol{V}=\boldsymbol{t}(\boldsymbol{p}\cdot\boldsymbol{I}-\boldsymbol{V})$ (from (3)), and relations (19) (or rather the corresponding relations for $\bar{G}$) show that the polynomials of the system represented by $(\boldsymbol{p}\cdot\boldsymbol{I}-\boldsymbol{V})\cdot\boldsymbol{x}$ belong to $\mathfrak{b}$; moreover, if $\boldsymbol{y}$ is a one-column matrix representing a system of polynomials of $\mathfrak{b}$, $\boldsymbol{M}\cdot\boldsymbol{y}$ is also such a matrix, for an arbitrary $n\times m$ matrix $\boldsymbol{M}$ with elements in $\mathcal{E}^{+}(K)$. Let us now prove the converse; we can write $\boldsymbol{C}=\Gamma_0+\boldsymbol{t}\Gamma_1+\cdots+\boldsymbol{t}^k\Gamma_k+\cdots$, where $\Gamma_k=(C_k^{(ij)})$ is an $n\times m$ matrix with elements in the ring $\mathfrak{W}(K)$ of the Witt-vectors. If we denote by $c_h^{(i)}(T_{01},\cdots,T_{hm})$ the polynomials of the system represented by $\boldsymbol{C}\cdot x$, and if $C_k^{(ij)}=(c_{k0}^{(ij)},c_{k1}^{(ij)},\cdots)$ $(c_{kh}^{(ij)}\varepsilon K)$, we have

$$c_0^{(i)}(T_{01},\cdots,T_{0m})=\sum_{j=1}^{m}c_{00}^{(ij)}T_{0j}.$$

Now we have seen in the proof of Theorem 2 (no. 6) that such a polynomial cannot belong to $\mathfrak{b}$ unless all the $c_{00}^{(ij)}=0$; but, as K is perfect, this means that there is a matrix Λ_0 with elements in $\mathfrak{W}(K)$, such that $\Gamma_0=\Lambda_0\boldsymbol{\pi}$. We can therefore write $\boldsymbol{C}=\Lambda_0(\boldsymbol{\pi}\cdot\boldsymbol{I}-\boldsymbol{t}\cdot\boldsymbol{V})+\boldsymbol{t}\cdot\boldsymbol{C}_1$, and $(\boldsymbol{t}\cdot\boldsymbol{C}_1)\cdot\boldsymbol{x}$ must belong to the ideal $\mathfrak{b}$; in other words, we are reduced to the case in which $\Gamma_0=0$. More generally, suppose Γ_k is the first coefficient in $\boldsymbol{C}$ which is $\neq 0$; we then have

$$c_k^{(i)}(T_{01},\cdots,T_{0m};\cdots;T_{k1},\cdots,T_{km})=\sum_{j=1}^{m}c_{k0}^{(ij)}T_{0j}$$

and the same argument as before shows that all the elements $c_{k0}^{(ij)}$ must be 0. By an easy inductive argument, it follows that $\boldsymbol{C}=\boldsymbol{L}(\boldsymbol{\pi}\cdot\boldsymbol{I}-\boldsymbol{t}\cdot\boldsymbol{V})$ with $\boldsymbol{L}=\Lambda_0+\boldsymbol{t}\Lambda_1+\cdots+\boldsymbol{t}^k\Lambda_k+\cdots$, the Λ_k being $n\times m$ matrices with elements in $\mathfrak{W}(K)$. This concludes the proof of the *necessity* of condition (32).

16. To prove that the condition is *sufficient*, we consider the homomorphism $\boldsymbol{r}$ of the algebra of polynomials P into the algebra of polynomials Q, defined by

$$\boldsymbol{r}(S_{hi})=w_h^{(i)}(T_{01},\cdots,T_{0m};\cdots;T_{h1},\cdots,T_{hm}) \tag{34}$$
$$(1\leq i\leq n;h=0,1,\cdots)$$

Condition (32) insures that the image of the ideal $\mathfrak{a}$ under the homomorphism $\boldsymbol{r}$ is contained in $\mathfrak{b}$; hence $\boldsymbol{r}$ defines a homomorphism $\boldsymbol{s}'$ of $P/\mathfrak{a}$ into $Q/\mathfrak{b}$, satisfying (31). It remains to be seen that this homomorphism is the derived homomorphism of a homomorphism $\boldsymbol{s}$ of G into $\bar{G}$; this will be done by showing that the conditions of [5, Prop. 2] are satisfied.

Introducing new sequences of indeterminates S'_{hi}, T'_{hi}, and using the notations of no. 6, we first remark that, as $\boldsymbol{r}$ is a homomorphism of P into Q, the image by $\boldsymbol{r}\otimes\boldsymbol{r}$ of the element

$$\sum_{0\leq\beta\leq\alpha} E_\beta(S_{01},\cdots,S_{hn},\cdots)E_{\alpha-\beta}(S'_{01},\cdots,S'_{hn},\cdots)$$

is obtained by replacing in that expression each S_{hi} (resp. S'_{hi}) by its value $\boldsymbol{r}(S_{hi})$ (resp. $\boldsymbol{r}(S'_{hi})$) given by (34). But from the fact that $\boldsymbol{w}$ is a homomorphism and the arguments in no. 8, it follows that we obtain in that way the expression

$$E_\alpha(\Psi_{01},\cdots,\Psi_{h1},\cdots;\cdots;\Psi_{0n},\cdots,\Psi_{hn})$$

where

$$\Psi_{hi}=w_h^{(i)}(T''_{01},\cdots,T''_{0m};\cdots;T''_{h1},\cdots,T''_{hm})$$

and

$$T''_{hi}=\phi_h(T_{0i},T'_{0i},T_{1i},T'_{1i},\cdots,T_{hi},T'_{hi})$$

($\phi=(\phi_h)$ now stands, as in no. 8, for the group law of the hyperexponential group H). We thus obtain a polynomial in the T''_{hi}, which (by the argument of no. 7) can be expressed as a linear combination of the polynomials

$$E_\lambda(T''_{01},\cdots,T''_{0m};\cdots;T''_{h1},\cdots,T''_{hm};\cdots). \tag{35}$$

Let us now go over to the quotient algebras $P/\mathfrak{a}$ and $Q/\mathfrak{b}$; let X'_{hi}, Y'_{hi} correspond to S'_{hi}, T'_{hi} by the natural mapping, and, changing slightly the notations of no. 8, let us write

$$Z_\alpha{}^0=\sum_{0\leq\beta\leq\alpha} E_\beta(X_{01},\cdots,X_{hn},\cdots)E_{\alpha-\beta}(X'_{01},\cdots,X'_{hn},\cdots)$$

$$\bar{Z}_\lambda{}^0=\sum_{0\leq\mu\leq\lambda} E_\mu(Y_{01},\cdots,Y_{hm},\cdots)E_{\lambda-\mu}(Y'_{01},\cdots,Y'_{hm},\cdots).$$

Then the image, by $\boldsymbol{s}'\otimes\boldsymbol{s}'$, of $Z_\alpha{}^0$, is obtained as a linear combination of the elements (35) in which T_{hi} and T'_{hi} have been respectively replaced by Y_{hi} and Y'_{hi}; but these elements are merely the elements $\bar{Z}_\lambda{}^0$, by the argument of no. 8. The conditions of [5, Prop. 2] are thus satisfied, and the proof of Theorem 4 is brought to an end.

17. We will say that the homomorphism $\boldsymbol{s}$ of G into $\bar{G}$, defined above, is *associated* to the matrix $\boldsymbol{W}$. We now prove (with the same notations):

THEOREM 5. *Suppose K is a perfect field. Every homomorphism $\boldsymbol{s}$ of G into $\bar{G}$ is associated to an $n\times m$ matrix $\boldsymbol{W}$ with elements in $\mathcal{E}^+(K)$.*

We are going to construct a matrix $\boldsymbol{W}$ for which relations (31) will hold. We can write first

$$\boldsymbol{s}'(X_{0i}) = \sum_{j=1}^{m} a_{ij} Y_{0j} \qquad (1 \leq i \leq n).$$

We consider the Witt-vectors $A_{ij} = (a_{ij}, 0, 0, \cdots)$, and the matrix $\boldsymbol{W}^{(0)} = (A_{ij})$; we have therefore

$$\boldsymbol{s}'(X_{0i}) = w_0^{(0,i)}(Y_{01}, \cdots, Y_{0m}) \qquad (1 \leq i \leq n).$$

As a basis for an inductive argument, suppose we have constructed a matrix $\boldsymbol{W}^{(r)}$ such that

$$\boldsymbol{s}'(X_{hi}) = w_h^{(r,i)}(Y_{01}, \cdots, Y_{0m}; \cdots; Y_{h1}, \cdots, Y_{hm}) \tag{36}$$

for $1 \leq i \leq n$ and $0 \leq h \leq r$. As $\boldsymbol{s}'$ is a homomorphism of hyperalgebras, we have, by (18)

$$\boldsymbol{s}'(Z_\alpha) = E_\alpha(\boldsymbol{s}'(X_{01}), \cdots, \boldsymbol{s}'(X_{hn}), \cdots). \tag{37}$$

We prove that the differences

$$U_i = \boldsymbol{s}'(X_{r+1,i}) - w_{r+1}^{(r,i)}(Y_{01}, \cdots, Y_{0m}; \cdots; Y_{r+1,1}, \cdots, Y_{r+1,m})$$

are *derivations* in $\bar{G}$ $(1 \leq i \leq n)$. We have by definition

$$\begin{aligned}(\boldsymbol{s}' \otimes \boldsymbol{s}')(X_{r+1,i}{}^0) &= I \otimes \boldsymbol{s}'(X_{r+1,i}) + \boldsymbol{s}'(X_{r+1,i}) \otimes I \\ &\quad + \sum_{0<k<p^{r+1}} \boldsymbol{s}'(Z_{k\epsilon_i}) \otimes \boldsymbol{s}'(Z_{(p^{r+1}-k)\epsilon_i}).\end{aligned} \tag{38}$$

Now in the sum in the right hand side of this formula, we can replace $\boldsymbol{s}'(Z_{k\epsilon_i})$ by its expression (37), and as $k < p^{r+1}$, only terms $\boldsymbol{s}'(X_{hi})$ with $h \leq r$ will intervene. These can then be replaced by their expression (36). Now, using (22) with $\alpha = p^{r+1}\epsilon_i$ (and therefore $E_\alpha(\boldsymbol{x}) = x_{r+1}{}^{(i)}$), we see that we have

$$\begin{aligned}&w_{r+1}^{(r,i)}(Y_{01}{}^0, \cdots, Y_{0m}{}^0; \cdots; Y_{r+1,1}{}^0, \cdots, Y_{r+1,m}{}^0) \\ &= I \otimes w_{r+1}^{(r,i)}(Y_{01}, \cdots, Y_{r+1,m}) + w_{r+1}^{(r,i)}(Y_{01}, \cdots, Y_{r+1,m}) \otimes I \\ &\quad + \sum_{0<k<p^{r+1}} \boldsymbol{s}'(Z_{k\epsilon_i}) \otimes \boldsymbol{s}'(Z_{(p^{r+1}-k)\epsilon_i})\end{aligned} \tag{39}$$

since $\boldsymbol{W}^{(r)}$ defines a homomorphism $\boldsymbol{w}^{(r)}$ of H^m into H^n. As computations with tensor products and with "Leibniz formulae" are "dual" to each other (see for instance [4, proof of Lemma 2]), comparison of (38) and (39) shows that $U_i(fg) = f \cdot U_i g + g \cdot U_i f$.

We can therefore write

$$U_i = \sum_{j=1}^{m} a_{ij}^{(r+1)} Y_{0j} \qquad (1 \leq i \leq n).$$

Consider the Witt vectors $A_{ij}^{(r+1)} = (a^{(r+1)}, 0, 0, \cdots)$, and the matrix $\boldsymbol{W}^{(r+1)} = \boldsymbol{W}^{(r)} + \boldsymbol{t}^{r+1} \cdot (A_{ij}^{(r+1)})$. It is then clear that we have the relations

$$\boldsymbol{s}'(X_{hi}) = w_h^{(r+1,i)}(Y_{01}, \cdots, Y_{0m}; \cdots; Y_{h1}, \cdots, Y_{hm})$$

for $1 \leq i \leq n$ and $0 \leq h \leq r+1$. The induction is then concluded in the usual way, the matrices $\boldsymbol{W}^{(r)}$ "converging" obviously towards a matrix $\boldsymbol{W}$ for which relations (31) are verified for every h. Theorem 5 is therefore proved.

18. The result embodied in Theorems 4 and 5 can be put in a much more suggestive form. Multiplying on the left both sides of (32) by $\boldsymbol{t}$ gives an equivalent condition, since the ring $\mathcal{E}^+(K)$ has no zero-divisors; using the commutation rules (7), we see therefore that (32) is equivalent to the relation

$$(\pi \boldsymbol{I}_n - \boldsymbol{t}\boldsymbol{U})\boldsymbol{W} = (\boldsymbol{W} + \boldsymbol{t}\boldsymbol{L})(\pi \boldsymbol{I}_m - \boldsymbol{t}\boldsymbol{V}) \tag{40}$$

($\boldsymbol{I}_m$ and $\boldsymbol{I}_n$ being the unit matrices of order m and n). Conversely, suppose there is an $n \times m$ matrix $\boldsymbol{W}_1$ such that

$$(\pi \boldsymbol{I}_n - \boldsymbol{t}\boldsymbol{U})\boldsymbol{W} = \boldsymbol{W}_1(\pi \boldsymbol{I}_m - \boldsymbol{t}\boldsymbol{V}). \tag{41}$$

From such a relation, it follows at once that $\boldsymbol{W}$ and $\boldsymbol{W}_1$ are congruent mod. $\boldsymbol{t}$, in other words we can write $\boldsymbol{W}_1 = \boldsymbol{W} + \boldsymbol{t}\boldsymbol{L}$, and we see therefore that (41) implies (40), hence also (32). The existence of an $n \times m$ matrix $\boldsymbol{W}_1$ over $\mathcal{E}^+(K)$ satisfying (41) is therefore (when K is perfect) the necessary and sufficient condition for $\boldsymbol{W}$ to represent a homomorphism of G into $\bar{G}$. But this can also be formulated in different terms: consider the left-modules $E = (\mathcal{E}^+(K))^n$, $F = (\mathcal{E}^+(K))^m$, direct sums of n (resp. m) modules identical to $\mathcal{E}^+(K)$ (considered as a left module over itself). In E (resp. F), let M (resp. N) be the submodule which is the image of E (resp. F) under the endomorphism represented by the matrix $\pi \boldsymbol{I}_n - \boldsymbol{t}\boldsymbol{U}$ (resp. $\pi \boldsymbol{I}_m - \boldsymbol{t}\boldsymbol{V}$). Relation (41) is then the necessary and sufficient condition for the matrix $\boldsymbol{W}$ to represent an $(\mathcal{E}^+(K))$-homomorphism of the module E into the module F, which sends M into N, and therefore to define *a homomorphism of E/M into F/N*. We thus see that, by associating with the group G the $\mathcal{E}^+(K)$-module E/M, *the theory of abelian formal Lie groups over a perfect field K is essentially reduced to the theory of left $\mathcal{E}^+(K)$-modules.*

19. Applications: I. Classification of abelian Lie groups. If $\boldsymbol{s}$ is an isomorphism of G onto $\bar{G}$ (for $m = n$), the matrix (a_{ij}) (with the notations of no. 17) must be invertible, and from that remark it is very easy to conclude that both matrices $\boldsymbol{W}$ and $\boldsymbol{W}_1$ in (41) must be invertible. Therefore:

THEOREM 6. *Let K be a perfect field. In order that two abelian Lie groups G, $\bar{G}$ over K, respectively associated to two square matrices $\boldsymbol{U}$, $\boldsymbol{V}$ of order n, with elements in $\mathcal{E}^{+}(K)$, be isomorphic, it is necessary and sufficient that $\pi\boldsymbol{I} - t\boldsymbol{U}$ and $\pi\boldsymbol{I} - t\boldsymbol{V}$ be equivalent.*

In the language of modules, this (with the notations of no. 18) means that the modules E/M and E/N associated to G and $\bar{G}$ are *isomorphic*. The classification of abelian Lie groups over a perfect field K is therefore equivalent to the classification of the $\mathcal{E}^{+}(K)$-modules having a finite system of generators to which they correspond; as mentioned in the Introduction, no satisfactory theory is known at present for such modules, due to the fact that $\mathcal{E}^{+}(K)$ is not a principal ideal ring.

It follows of course from (32) that G and $\bar{G}$ will be isomorphic if there exists an invertible matrix $\boldsymbol{W}$ such that $\boldsymbol{W}^{\sigma}\boldsymbol{U}\boldsymbol{W}^{-1} = \boldsymbol{V}$, but this condition is by no means necessary, as we shall see below, and the problem of classifying matrices under this stronger equivalence relation does not appear to be much easier, even when K is algebraically closed. When $\boldsymbol{U}$ is invertible and K algebraically closed, there is always an invertible matrix $\boldsymbol{W}$ such that $\boldsymbol{W}^{\sigma}\boldsymbol{U}\boldsymbol{W}^{-1} = \boldsymbol{I}$, as was essentially proved in [4, no. 13]; the corresponding groups are therefore all isomorphic to $(W_1^{*})^n$. A similar proof would show that, when K is algebraically closed and $\boldsymbol{U} = \boldsymbol{t}^r\boldsymbol{U}_1$, with $\boldsymbol{U}_1$ invertible, there exists $\boldsymbol{W}$ such that $\boldsymbol{W}^{\sigma}\boldsymbol{U}\boldsymbol{W}^{-1} = \boldsymbol{t}^r\boldsymbol{I}$, and the corresponding groups are therefore isomorphic to the direct product $(I_r)^n$.

In particular, for the groups of "symmetric braid" type [5, no. 10], we have

$$\boldsymbol{U} = \begin{pmatrix} 0 & \boldsymbol{t}^r \\ \boldsymbol{t}^r & 0 \end{pmatrix} = \boldsymbol{t}^r \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

and therefore (as announced in [5]) if K is algebraically closed, these groups are isomorphic to products $I_r \times I_r$. Similarly, the groups of the "tree" type, corresponding to matrices

$$\begin{pmatrix} 0 & \boldsymbol{t}^r \\ 0 & \boldsymbol{t}^s \end{pmatrix}$$

do not yield new types of groups for $r \geq s$, for the matrix is then

$$\boldsymbol{U} = \boldsymbol{t}^s \begin{pmatrix} 0 & \boldsymbol{t}^{r-s} \\ 0 & I \end{pmatrix}$$

and it is readily verified that the matrix $\boldsymbol{W} = \begin{pmatrix} I & -\boldsymbol{t}^{r-s} \\ 0 & I \end{pmatrix}$ is such that $\boldsymbol{W}^{\sigma}\boldsymbol{U}\boldsymbol{W}^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & \boldsymbol{t}^s \end{pmatrix}$, so that the group is isomorphic to $W_1 \times I_s$ (as was conjectured in [5, no. 10]).

Going now to the other extreme, let us consider the case in which K is the prime field $\boldsymbol{F}_p$, and let us investigate the classification of *one-dimensional* abelian groups over K. As σ is then the identity, $\mathcal{E}^{+}(K)$ is a commutative ring, and the problems amounts, according to Theorem 6, to classify elements of that ring under the equivalence relation

$$\boldsymbol{\pi} - \boldsymbol{t}\boldsymbol{a} = \boldsymbol{c}(\boldsymbol{\pi} - \boldsymbol{t}\boldsymbol{b}) \tag{42}$$

where $\boldsymbol{c}$ is invertible. We have of course

$$\boldsymbol{a} = A_0 + \boldsymbol{t}A_1 + \cdots + \boldsymbol{t}^k A_k + \cdots$$

where the $A_k = (a_{k0}, a_{k1}, \cdots)$ are Witt-vectors with components in $\boldsymbol{F}_p$, and similarly for $\boldsymbol{b}$ and $\boldsymbol{c}$. It is then easily verified that, given $\boldsymbol{b}$, it is possible to determine $\boldsymbol{c}$ in one and only one way, such that all the elements $a_{ki} = 0$ for $i > 0$. The "first" elements a_{k0} are then invariants of the class of abelian groups; as they can be chosen arbitrarily in $\boldsymbol{F}_p$, we see that the set of classes of non isomorphic abelian groups (isomorphism being meant in the sense of $\boldsymbol{F}_p$-isomorphism) has the *power of the continuum* (in sharp distinction to the corresponding situation in the case of an algebraically closed field K). We may also observe that in this case, the relation $\boldsymbol{W}^{\sigma}\boldsymbol{U}\boldsymbol{W}^{-1} = \boldsymbol{V}$ would reduce to $\boldsymbol{a} = \boldsymbol{b}$ instead of (42).

20. Applications: II. Analytically simple abelian Lie groups. We now want to show that there exist abelian Lie groups of dimension > 1 which have *no nontrivial subgroup*, and can therefore be called "analytically simple" groups. More precisely, we will exhibit a group G of dimension 2 such that there exists *no non zero homomorphism of a group of dimension* 1 *into* G. G will be a group of the "dissymetric braid" type, corresponding to the matrix $\begin{pmatrix} 0 & \boldsymbol{t}^2 \\ \boldsymbol{t} & 0 \end{pmatrix}$. We may suppose that K is algebraically closed, and then a one-dimensional group will correspond to a one-element matrix of type $\boldsymbol{t}^q$. The existence of a homomorphism of this last group into G means that there exist elements $\boldsymbol{a}$, $\boldsymbol{a}'$, $\boldsymbol{b}$, $\boldsymbol{b}'$ in $\mathcal{E}^{+}(K)$, such that

$$(\boldsymbol{\pi} - \boldsymbol{t}^{q+1})(\boldsymbol{a}\ \ \boldsymbol{a}') = (\boldsymbol{b}\ \ \boldsymbol{b}')\begin{pmatrix} \boldsymbol{\pi} & -\boldsymbol{t}^3 \\ -\boldsymbol{t}^2 & \boldsymbol{\pi} \end{pmatrix} \tag{43}$$

which is equivalent to

$$\begin{cases} (\boldsymbol{\pi} - \boldsymbol{t}^{q+1})\boldsymbol{a} = \boldsymbol{b}\boldsymbol{\pi} - \boldsymbol{b}'\boldsymbol{t}^2 \\ (\boldsymbol{\pi} - \boldsymbol{t}^{q+1})\boldsymbol{a}' = \boldsymbol{b}'\boldsymbol{\pi} - \boldsymbol{b}\boldsymbol{t}^3. \end{cases} \tag{44}$$

We write $\boldsymbol{a} = \sum_k \boldsymbol{t}^k A_k$, $\boldsymbol{a}' = \sum_k \boldsymbol{t}^k A'_k$, $\boldsymbol{b} = \sum_k \boldsymbol{t}^k B_k$, $\boldsymbol{b}' = \sum_k \boldsymbol{t}^k B'_k$, where the

A_k, A'_k, B_k, B'_k are Witt vectors. It follows from the proof of Theorem 5 (or by a direct argument from (41)) that we can suppose $A_k=(a_{k0},0,0,\cdots)$, $A'_k=(a'_{k0},0,0,\cdots)$; on the other hand, we can suppose that $\boldsymbol{b}$ and $\boldsymbol{b}'$ are not both divisible by a power of $\boldsymbol{t}$, since otherwise $\boldsymbol{a}$ and $\boldsymbol{a}'$ would also be divisible by it, and we would then be reduced to the first case. We remark finally that we have $q>0$, since there are no homomorphisms of the multiplicative group into G which are not trivial, the hyperalgebra of G being a nilalgebra.

1. Suppose $B_0=(b_{00},b_{01},\cdots)\neq 0$. Identifying the coefficients of $\boldsymbol{t}^0$ in the first equation (44) gives $A_0=B_0$, hence $a_{00}=b_{00}\neq 0$ and $b_{0i}=0$ for $i>0$. We cannot have $q=1$, for in that case, identifying the coefficients of $\boldsymbol{t}^2$ in the second equation (44) would give $\pi A'_2-A'_0=\pi B'_2$, hence $a'_{00}=0$, and $A'_0=0$; but then, the same equation would give $B'_0=A'_0=0$, and identifying the coefficients of $\boldsymbol{t}^2$ in the first equation (44) would give $\pi A_2-A_0=\pi B_2$, which contradicts $a_{00}\neq 0$. The identification of the coefficients of $\boldsymbol{t}^2$ in the first equation gives then (since $q>1$), $\pi A_2=\pi B_2-B'^{\sigma^2}_0$, and if $B'_0=(b'_{00},b'_{01},\cdots)$, this yields $b'_{00}=0$. As $A'_0=B'_0$, we have also $a'_{00}=0$. Now, if $q=2$, we obtain by identifying the terms in $\boldsymbol{t}^3$ in the second equation (44),

$$\pi A'_3-A'_0=\pi B'_3-B_0^{\sigma^3}$$

and this yields $b_{00}=0$, contrary to assumption; if on the other hand $q>2$, we obtain similarly $\pi A'_3=\pi B'_3-B_0^{\sigma^3}$, and we reach again a contradiction.

2. Suppose now $B'_0\neq 0$; then we have in the same way $A'_0=B'_0$, hence $a'_{00}=b'_{00}\neq 0$ and $b'_{0i}=0$ for $i>0$. We see as in 1. that we cannot have $q=1$: but if $q>1$, the relation $\pi A_2=\pi B_2-B'^{\sigma^2}$ gives $b'_{00}=0$, hence again a contradiction. It is therefore impossible to satisfy (43) when $\boldsymbol{a}$ and $\boldsymbol{a}'$ are not both 0.

A similar argument shows that if G is associated to the matrix $\begin{pmatrix}0 & \boldsymbol{t}^s\\ \boldsymbol{t}^r & 0\end{pmatrix}$ with $r<s$, then G is analytically simple if $s+r$ is *odd*; but if $s+r=2h$, there is a nontrivial homomorphism of the one-dimensional group corresponding to $\boldsymbol{t}^h$ into G. These examples seem to show that the determination of all abelian analytically simple groups is a problem of great complexity.

21. Applications: III. Duality of abelian Lie groups. Suppose that K is the *prime field* $\boldsymbol{F}_p$; then the ring $\mathscr{E}^+(K)$ is commutative and the transposed matrix of any matrix with elements in that ring is again a matrix with elements in $\mathcal{E}^{+}(K)$ (whereas in general it is a matrix over the opposite ring). If, to each square matrix $\boldsymbol{U}$ with elements in $\mathcal{E}^{+}(K)$, we associate its transposed matrix ${}^{t}\boldsymbol{U}$, we are defining a pairing between the associated groups, which we can call *dual* to each other. If the matrix $\boldsymbol{W}$ represents a homomorphism of G into $\bar{G}$, then it follows from (41) that ${}^{t}\boldsymbol{W}_1$ represents a homomorphism of the dual $\bar{G}^{*}$ of $\bar{G}$ into the dual G^{*} of G; we will say that this homomorphism is *transposed* of the homomorphism represented by $\boldsymbol{W}$; the usual rules concerning transposed homomorphisms are then valid.

One-dimensional groups (over $\boldsymbol{F}_p$) are obviously isomorphic to their dual; so are two-dimensional groups of the "streak" or "braid" type; it is an open question whether an abelian Lie group is always isomorphic to its dual.

NORTHWESTERN UNIVERSITY.

BIBLIOGRAPHY.

[1] J. Dieudonné, "Groupes de Lie et hyperalgèbres de Lie sur un corps de caractéristique $p > 0$," *Commentarii Mathematici Helvetici*, vol. 28 (1954), pp. 87-118.

[2] ———, "Sur la notion de variables canoniques," *Anais da Academia Brasileira de Ciencias*, vol. 27 (1955).

[3] ———, "Le calcul différentiel dans les corps de caractéristique $p > 0$," *Proceedings of the International Congress of Mathematicians*, 1954, vol. 2.

[4] ———, "Lie groups and Lie hyperalgebras over a field of characteristic $p > 0$, (II)," *American Journal of Mathematics*, vol. 77 (1955), pp. 218-244.

[5] ———, "Groupes de Lie et hyperalgèbres de Lie sur un corps de caractéristique $p > 0$ (III)," *Mathematische Zeitschrift*, vol. 62 (1955).

[6] ———, "Witt groups and hyperexponential groups," *Mathematika*, vol. 2 (1955).

[7] E. Witt, "Zyklische Körper und Algebren der Charakteristik p vom Grad p^n," *Journal für die Reine und Angewandte Mathematik*, vol. 176 (1937), pp. 126-140.

ON THE LOCAL BEHAVIOR OF SOLUTIONS OF NON-PARABOLIC PARTIAL DIFFERENTIAL EQUATIONS.

III. Approximations by spherical harmonics.*

By PHILIP HARTMAN and AUREL WINTNER.

In this paper, the analogues of the theorems of [4], [5] on solutions of elliptic partial differential equations will be obtained for the case where the number of independent variables exceeds 2. For the sake of notational simplicity, it will be assumed that the number of independent variables is 3. The equation to be considered is of the type $\Delta u + \cdots = 0$, where Δu is the Euclidean Laplacian of u, and no second order partial derivative of u occurs in the rest of the equation. The replacement of Δu by a more general linear combination of second derivatives of u will not be considered at this time. (In the plane, this more general case can be reduced to the special case by conformal mappings under suitable smoothness assumptions on the coefficients; in space, perturbation methods of Korn and Lichtenstein can be used.) For simplicity, the partial differential equation to be considered will be assumed to be linear (the methods are applicable to non-linear equations of the type (5)-(6) below).

The first part of the paper deals with solutions near a zero, the second part with solutions near an isolated singularity.

Part 1. The zeros of a solution.

1. Statement of the theorems. In the linear elliptic partial differential equation

$$\Delta u + a \cdot \nabla u + \gamma u = 0, \tag{1}$$

let Δu denote the (Euclidean) Laplacian, ∇u the gradient of u, $a = a(x)$ a vector and $\gamma = \gamma(x)$ a scalar function of the position vector x, finally $a \cdot \nabla u$ the scalar product of a and ∇u. It is known ([4], pp. 449-450) that if the number of independent variables in (1) is 2, say $x = (x_1, x_2)$, if $a(x)$ and $\gamma(x)$ are continuous in a vicinity of $x = 0$ and if $u = u(x)$ is a C^2-solution (or even a C^1-solution in an appropriate sense) vanishing at

---

* Received April 18, 1955.

$x=0$, then either $u(x)\equiv 0$ or there exist an integer $N>0$ and constants $(c_1,c_2)\neq(0,0)$ such that, as $r=|x|\rightarrow 0$,

$$\begin{aligned} u(x)&=c_1r^N\cos N\phi+c_2r^N\sin N\phi+o(r^N),\\ \nabla u(x)&=\nabla(c_1r^N\cos N\phi+c_2r^N\sin N\phi)+o(r^{N-1}), \end{aligned} \tag{2}$$

uniformly in ϕ, where $x=(x_1,x_2)=(r\cos\phi, r\sin\phi)$. The proof of this fact depended on an adaptation of a proof of a uniqueness theorem of Carleman [2].

The object of this chapter is to obtain similar asymptotic formulae in case the number of independent variables in (1) exceeds 2. The methods used to prove (2), depending on function theory, cannot be applied in this situation. They will be replaced by procedures depending on Fourier analysis and on a generalization of the argument used by C. Müller [6] to obtain an extension of Carleman's uniqueness theorem in the plane to a uniqueness theorem in space.

The analogue of (2) will be

$$\begin{aligned} u(x)&=r^NS_N(e)+o(r^N),\\ \nabla u(x)&=\nabla(r^NS_N(e))+o(r^{N-1}), \end{aligned} \tag{3}$$

where $S_N(e)$ is a spherical harmonic of order N. Here and in the sequel, $x=re$, where $r=|x|$; so that e is a unit vector.

THEOREM 1. *Let x be a 3-dimensional vector, $\alpha=\alpha(x)$ a 3-dimensional continuous vector and $\gamma=\gamma(x)$ a continuous scalar function in a sphere about $x=0$, say $0\leqq|x|\leqq 1$. Let $u=u(x)$ be a C^1-solution of (1) on $|x|\leqq 1$ satisfying*

$$u(0)=0. \tag{4}$$

Then either $u(x)\equiv 0$ on $|x|\leqq 1$ or there exist an integer $N>0$ and a spherical harmonic $S_N(e)\not\equiv 0$ of order N on the unit sphere $|x|=1$ with the property that the relations (3) hold uniformly in e, as $r=|x|\rightarrow 0$.

A function $u=u(x)$ is called a C^1-solution of (1) on $|x|\leqq 1$ if it is of class C^1 and can be represented as the sum of an harmonic function and the Newtonian potential of $(\alpha\cdot\nabla u+\gamma u)/4\pi$.

Remark 1. It will be clear from the proof that Theorem 1 remains correct if the dimension number 3 is replaced by any dimension number.

Remark 2. It will also be clear that Theorem 1 remains true if (1) is replaced by a non-linear equation

$$(5) \qquad \Delta u + g(x, u, \nabla u) = 0,$$

where $g(x, u, \nabla u)$ is a continuous function of its (seven) variables on a set $|x| \leqq 1$, $|u| \leqq \text{const.}$, $|\nabla u| \leqq \text{const.}$ and is subject to an inequality of the form

$$(6) \qquad |g(x, u, \nabla u)| \leqq \text{Const.}\ (|u| + |\nabla u|).$$

Remark 3. The $o(r^N)$- and $o(r^{N-1})$-terms in (3) can be replaced by $O(r^{N+1}|\log r|)$ and $O(r^N|\log r|)$, respectively. This improvement of (3) follows by repeating the arguments of Sections 6 and 7 below with (61) replaced by the estimates, $u = O(r^N)$ and $\nabla u = O(r^{N-1})$, supplied by (3). A similar remark applies to Theorems 1 bis and 4 below and to the corresponding theorems of [4], [5].

As a consequence of Theorem 1, there results the following corollary which is an analogue (and extension) of Carleman's theorem [2] for two independent variables:

COROLLARY 1. *Let* $\alpha(x)$, $\gamma(x)$ *and* $u(x)$ *satisfy the condition of Theorem* 1 *and, in addition, let*

$$(7) \qquad u(x) = o(r^n) \ \textit{for}\ n = 0, 1, \cdots$$

as $r \to 0$. *Then* $u(x) \equiv 0$ *on* $|x| \leqq 1$.

Another consequence of Theorem 1 is the following:

COROLLARY 2. *Let* $\alpha(x)$, $\gamma(x)$, $u(x)$ *satisfy the assumptions of Theorem* 1. *Then the zeros of* $\nabla u(x)$ *cannot cluster at* $x = 0$ *unless* $u(x) \equiv 0$ *on* $|x| \leqq 1$.

In another direction, Theorem 1 can be extended as follows:

THEOREM 1 bis. *Let* $\alpha = \alpha(x)$ *and* $\gamma = \gamma(x)$ *be, respectively, a vector and scalar function of class* C^1 *in a vicinity of* $x = 0$. *Let* $u = u(x) \not\equiv 0$ *be a* C^2*-solution of* (1) *in a vicinity of* $x = 0$ *satisfying* (4) (*so that there exist an integer* $N > 0$ *and a spherical harmonic* $S_N(e) \not\equiv 0$ *satisfying* (3)). *Then the Hessian matrix of* u, *say* u_{xx}, *satisfies, as* $r \to 0$,

$$(3\,\text{bis}) \qquad u_{xx} = (r^N S_N(e))_{xx} + o(r^{N-2}).$$

The details of the proof of this theorem will be omitted. It can be obtained by considering the inhomogeneous system of equations,

$$\Delta u_k + \alpha \cdot \nabla u_k + \gamma u_k + \alpha_k \cdot \nabla u + \gamma_k u = 0, \qquad (k = 1, 2, 3),$$

for the components of the gradient $\nabla u = (u_1, u_2, u_3)$, obtained by differentiating (1), and modifying the proof of Theorem 1 in a manner analogous to that used in [4], pp. 467-469.

2. An estimate for $\| u - u^N \|$.

In what follows, x will denote a 3-dimensional vector. When it is necessary to deal with components of x, these will be denoted by (x, y, z); there will be no confusion over this double use of x. The geographical coordinates of a point e on the unit sphere

$$B: \ |e| = 1 \tag{8}$$

will be denoted by (θ, ϕ). Hence $x = (r \cos\phi \sin\theta, r \sin\phi \sin\theta, r \cos\theta)$, where $r = |x|$. A function $u = u(x)$ can be considered as a function of r and e, say $u = u(r, e)$. The symbol $\| \cdots \|$ will refer to the L^2-norm on (8); thus $\| u(r, e) \|$ becomes a function of r. The element of area on (8) will be denoted by $d\sigma$.

Let $S_{nj}(e)$ be the spherical harmonic

$$S_{nj} = c_{nj} P_{nj}(\theta) \cos j\phi \text{ or } S_{n\, j+n} = c_{nj} P_{nj}(\theta) \sin j\phi \tag{9}$$

for $j = 0, 1, \cdots, n$ or $j = 1, \cdots, n$, respectively, and $n = 0, 1, \cdots$, where the $P_{nj}(\theta)$ are the Legendre functions and the c_{nj} are normalizing factors. Thus S_{nj}, where $j = 0, \cdots, 2n$ and $n = 0, 1, \cdots$, is a complete orthonormal sequence on (8).

Let a function $u = u(x)$ of $x = re$, where $r = |x|$, be written as a function $u(r, e)$. If, for a fixed $r > 0$, the functions $f(x) = f(r, e)$ and $u(x) = u(r, e)$ are of class L^2 on (8), let

$$(10) \quad F_{nj}(r) = \iint_B f(r, e) S_{nj}(e) d\sigma, \qquad (11) \quad U_{nj}(r) = \iint_B u(r, e) S_{nj}(e) d\sigma$$

denote the respective Fourier constants. Thus, $u(r, e)$ has the Fourier series

$$u(r, e) \sim \sum_{n=0}^{\infty} \sum_{j=0}^{2n} U_{nj}(r) S_{nj}(e); \tag{12}$$

$u^N(r, e)$ will denote the partial sum

$$u^N(r, e) = \sum_{n=0}^{N} \sum_{j=0}^{2n} U_{nj}(r) S_{nj}(e). \tag{13}$$

Let $f(x)$ be a continuous function on $|x| \leqq 1$ and let $u = u(x)$ be a C^1-solution of the Poisson equation

$$\Delta u = f. \tag{14}$$

The main equalities will furnish estimates, as $r\to 0$, for $U_{nj}(r)$ and $\| u(r,e)-u^N(r,e)\|^2$ in terms of

$$(15)\qquad \Phi(r)=\| f(r,e)\|^2.$$

LEMMA 1. *Let* $f(x)$ *be a continuous function on* $|x|\leqq 1$ *and let* $u(x)$ *be a* C^1*-solution of* (14). *Let* λ *be a (non-negative) number and* N *an integer satisfying*

$$(16)\qquad N>\lambda+\tfrac{1}{2}.$$

Then there exists a constant C *(independent of* N *and* λ*) with the property that, for* $0<r\leqq 1$,

$$(17)\qquad \| u-u^N\|^2\leqq C\chi(r),$$

where

$$(18)\qquad \chi(r)=r^{2N+2}+r^{2\lambda+3}(2N-2\lambda-1)^{-1}\int_0^1 s^{-2\lambda}\Phi(s)\,ds\quad(\leqq\infty).$$

The constant C depends only on bounds for $|u|$, $|\nabla u|$ on $|x|=1$.

3. Proof of Lemma 1. If $u=u(x)$ is of class C^2, then differentiations under the integral sign show that, by virtue of (14) and (10), the function (11) satisfies the singular inhomogeneous differential equation

$$(19)\qquad (r^2U')'-n(n+1)U=r^2F,\qquad\text{where } '=d/dr,$$

$U=U_{nj}$ and $F=F_{nj}$. The general solution $U=U(r)$ of (19), continuous at $r=0$, is

$$(20)\qquad U(r)=(2n+1)^{-1}\{r^n(A-\int_r^1 s^{1-n}F(s)ds)-r^{-1-n}\int_0^r s^{n+2}F(s)ds\},$$

where $A=A_{nj}$ is a constant. Since u is of class C^1, it follows that the derivative

$$(21)\qquad U_{nj}'(r)=\iint_B u_r(r,e)S_{nj}(e)\,d\sigma$$

exists, is continuous and can be obtained by a differentiation of (20):

$$(22)\ U'(r)=(2n+1)^{-1}\{nr^{n-1}(A-\int_r^1 s^{1-n}F(s)ds)+(n+1)r^{-2-n}\int_0^r s^{n+2}F(s)ds\}.$$

The formulae (20) and (22), just derived under the assumption that

$u=u(x)$ is of class C^2, are valid if $u=u(x)$ is only of class C^1. This follows, for example, by approximating f and u by smooth functions.

By letting $r=1$ in (20) and (22),

$$(23) \qquad A=(n+1)U(1)+U'(1) \qquad (A=A_{nj},\, U=U_{nj}).$$

Hence, by Bessel's inequality,

$$(24) \qquad \sum_{n=0}^{\infty}\sum_{j=0}^{2n} A^2/(n+1)^2 \leqq \text{const.}\,(\|u\|^2+\|u_r\|^2)_{r=1}<\infty.$$

The relation (20) shows that

$$(25) \qquad \tfrac{1}{3}(2n+1)^2U^2 \leqq A^2r^{2n}+r^{2n}\Big(\int_r^1 s^{1-n}F(s)ds\Big)^2+r^{-2-2n}\Big(\int_0^r s^{n+2}F(s)ds\Big)^2.$$

In view of Schwarz's inequality,

$$\Big(\int_r^1 s^{1-n}F(s)ds\Big)^2 \leqq \Big(\int_r^\infty s^{2(1-n+\lambda)}ds\Big)\Big(\int_r^1 s^{-2\lambda}F^2(s)ds\Big),$$

if $2(n-\lambda-1)>1$, that is, if $n>\lambda+\frac{3}{2}$. Similarly,

$$\Big(\int_0^r s^{n+2}F(s)ds\Big)^2 \leqq \Big(\int_0^r s^{2(n+\lambda+2)}ds\Big)\Big(\int_0^r s^{-2\lambda}F^2(s)ds\Big).$$

The last three formula lines give

$$(26) \qquad \tfrac{1}{3}(2n+1)^2U^2 \leqq A^2r^{2n}+r^{2\lambda+3}(2n-2\lambda-3)^{-1}\int_r^1 s^{-2\lambda}F^2(s)ds$$
$$+\,r^{2\lambda+3}(2n+2\lambda+5)^{-1}\int_0^r s^{-2\lambda}F^2(s)ds,$$

if $n>\lambda+\frac{3}{2}$. If this relation is divided by $(2n+1)^2$ and the resulting inequality summed over $n=N+1, N+2,\cdots$ and $j=0,\cdots,2n$, where $N>\lambda+\frac{1}{2}$, then Parseval's relation shows that

$$(27) \qquad \|u-u^N\|^2 \leqq 3\sum_{n=N+1}^{\infty}\sum_{j=0}^{2n} r^{2n}A_{nj}^2/(2n+1)^2$$
$$+\,3r^{2\lambda+3}(2N-2\lambda-1)^{-1}\int_0^1 s^{-2\lambda}\Phi(s)ds.$$

In view of (24), this leads to (17)-(18).

4. The derivatives u_r, u_ϕ. Note that, in the deduction of (27) from (26), the factor $(2n+1)^{-2}$ in the expressions leading to the last term in (27) was replaced by 1. Since (22) and (25) show that

$$\tfrac{1}{3}(2n+1)^2 r^2 U'^2 \leqq (n+1)^2\{\cdot\ \cdot\ \cdot\}, \tag{28}$$

where $\{\cdot\ \cdot\ \cdot\}$ is the expression on the right side of the inequality (25), the proof of Lemma 1 implies

LEMMA 2. *Let u, f, N, λ satisfy the conditions of Lemma* 1. *Then there exist a constant C (independent of N and λ) with the properties that, for $0 < r < 1$,*

$$r^2 \| u_r - u_r^N \|^2 \leqq C\chi^*(r), \tag{29}$$

where

$$\chi^*(r) = N^2 r^{2N+2}(1-r^2)^{-3} + r^{2\lambda+3}(2N-2\lambda-1)^{-1}\int_0^1 s^{-2\lambda}\Phi(s)ds. \tag{30}$$

In fact, the analogue of (27) is

$$r^2 \| u_r - u_r^N \|^2 \leqq 3 \sum_{n=N+1}^{\infty} \sum_{j=0}^{2n} r^{2n} A_{nj}^2 + 3r^{2\lambda+3}(2N-2\lambda-1)^{-1}\int_0^1 s^{-2\lambda}\Phi(s)ds,$$

since $(n+1)^2/(2n+1)^2 \leqq 1$. In view of (24),

$$\sum_{j=0}^{2n} A_{nj}^2 = o(n^2) \quad \text{as} \quad n \to \infty.$$

Clearly, (29)-(30) follows from the last two formula lines.

A statement analogous to Lemma 2 will be obtained for the case where ru_r is replaced by the partial derivative u_ϕ.

LEMMA 3. *Let u, f, N, λ satisfy the conditions of Lemma* 1. *Then there exists a constant C (independent of N and λ) with the properties that, for $0 < r < 1$,*

$$\| u_\phi(r, e) - u_\phi^N(r, e) \|^2 \leqq C\chi^*(r). \tag{31}$$

In order to see this, let V_{nj} denote the Fourier constant

$$V_{nj}(r) = \iint_B u_\phi(r, e) S_{nj}(e)\, d\sigma \tag{32}$$

of u_ϕ. Since $d\sigma = \sin\theta\, d\theta d\phi$, a partial integration shows that

$$V_{nj}(r) = -\iint_B u(r, e) S_{nj\,\phi}(e)\, d\sigma. \tag{33}$$

It is clear from (9) that $S_{nj\,\phi} = -jS_{n\,j+n}$ if $j = 0, 1, \cdots, n$ and that $S_{nj\,\phi} = (j-n)S_{n\,j-n}$ if $j = n+1, \cdots, 2n$. Thus

$$(34) \qquad V_{nj} = -jU_{n\,j+n} \text{ or } V_{nj} = (j-n)U_{n\,j-n}$$

according as $0 \leqq j \leqq n$ or $n+1 \leqq j \leqq 2n$. The factor j or $j-n$, which is majorized by n, is immaterial (as was the factor $(n+1)^2$ on the right of (28)); so that the proofs of Lemmas 1 and 2 imply Lemma 3.

5. The principal lemmas. Let N be an integer and l a number (not necessarily an integer) subject to the inequalities

$$(35) \qquad 2 \leqq N \leqq l+2.$$

The estimates to follow will be based on the assumptions that, as $r \rightarrow 0$,

$$(36) \quad \| u(r,e) \| = o(r^{N-1}), \qquad (37) \quad \Phi(r) = o(r^{2l}).$$

These hypotheses, Schwarz's inequality, (10), (11) and (15) show that, as $r \rightarrow 0$,

$$(38) \quad | F_{nj}(r) | = o(r^{l}), \qquad (39) \quad | U_{nj}(r) | = o(r^{N-1}).$$

Since the last part of (20) is

$$(40) \qquad r^{-1-n} \int_0^r s^{n+2} F(s)\,ds = o(r^{l+2}) \text{ as } r \rightarrow 0,$$

it follows from $l+2 \geqq N$, (39) and (20) that, as $r \rightarrow 0$,

$$(41) \quad I = o(1) \text{ if } n < N, \text{ where } I = I_{nj}(r) = A - \int_r^1 s^{1-n} F(s)\,ds.$$

Thus, if $n < N$,

$$(42) \quad (2n+1)U_{nj}(r) = r^n \int_0^r s^{1-n}F(s)\,ds - r^{-1-n} \int_0^r s^{n+2} F(s)\,ds,$$

where the (conditional) convergence of the first integral is assured by (41).

Schwarz's inequality, as used in obtaining (26) from (25), shows that

$$(43) \qquad (2n+1)^2 U^2 \leqq (3 - 2n + 2\mu)^{-1} r^{2\mu+3} \int_0^r s^{-2\mu} F^2(s)\,ds,$$

if (42) holds (for example, if $n < N$) and if $n < \mu + \frac{3}{2}$. Omitting the factor $(2n+1)^2 \geqq 1$ on the left and summing over $n = 0, 1, \cdots, N-1$ and $j = 0, 1, \cdots, 2n$, it follows that, if $N - 1 < \mu + \frac{3}{2}$,

$$(44_1) \qquad \| u^{N-1} \|^2 \leqq (5 - 2N + 2\mu)^{-1} r^{2\mu+3} \int_0^r s^{-2\mu} \Phi(s)\, ds.$$

Similarly,

$$(44_2) \qquad r^2 \| u_r^{N-1} \|^2 \leqq (5 - 2N + 2\mu)^{-1} r^{2\mu+3} \int_0^r s^{-2\mu} \Phi(s)\, ds,$$

$$(44_3) \qquad \| u_\phi^{N-1} \|^2 \leqq (5 - 2N + 2\mu)^{-1} r^{2\mu+3} \int_0^r s^{-2\mu} \Phi(s)\, ds.$$

For later reference, this result will be stated as a lemma:

LEMMA 4. *Let the conditions on u, f in Lemma 1 hold. In addition, let there exist an integer N and numbers l, μ satisfying* (35)-(37) *and*

$$(45) \qquad N - 1 < \mu + \tfrac{3}{2} < l + 2.$$

Then (44_1)-(44_3) *hold.*

It is clear that the convergence of the integral in (44) is implied by (37) and the last part of (45).

It will be convenient to have also the following results on u^N:

LEMMA 5. *Let u, f, l and N satisfy the conditions of Lemma 4. Then, as $r \rightarrow 0$,*

$$(46) \qquad \| u^N \| = o(r^{N-\epsilon}) \text{ if } \epsilon > 0.$$

If, in addition, it is supposed that (35) *is replaced by*

$$(47) \qquad 2 \leqq N < l + 2,$$

then (46) *can be improved to*

$$(48) \qquad \| u^N - r^N S_N \| = o(r^N),$$

where $S_N = S_N(e)$ is some spherical harmonic of order N.

Since $N \leqq l + 2$, it follows that if $n = N$, then the expression $I = I(r)$ in (41) is $O(|\log r|) = o(r^\epsilon)$ as $r \rightarrow 0$, for every $\epsilon > 0$. Hence, by (20) and (40), $U_{nj} = o(r^{n-\epsilon})$ if $n = N$. In view of (44_1), this proves (46).

The last estimate for U_{nj} can be improved to the statement that

$$(49) \qquad \lim_{r \rightarrow 0} r^{-n} U_{nj}(r) = B_{nj} \text{ exists if } n = N,$$

provided that $\lim I(r)$ exists (when $n = N$). This is the case, for example, if (47) holds. Hence, if $S_N(e) = \sum_{j=0}^{2N} B_{Nj} S_{Nj}(e)$, then (48) follows from (44_1).

This verification of Lemma 5 and the arguments leading to Lemmas 2 and 3 have the following consequences:

LEMMA 6. *Let the assumptions of Lemma 4 hold. Then, as* $r \rightarrow 0$,

$$\| u_\phi{}^N \| = o(r^{N-\epsilon}) \text{ if } \epsilon > 0. \tag{50_1}$$

If, in addition, (47) *holds, then*

$$\| u_\phi{}^N - (r^N S_N)_\phi \| = o(r^N). \tag{50_2}$$

LEMMA 7. *Let the assumptions of Lemma 4 hold. Then, as* $r \rightarrow 0$,

$$r \| u_r{}^N \| = o(r^{N-\epsilon}) \textit{ if } \epsilon > 0. \tag{51_1}$$

If, in addition, (47) *holds, then*

$$r \| u_r{}^N - (r^N S_N)_r \| = o(r^N). \tag{51_2}$$

6. $\| \boldsymbol{u} \|$ **and** $\| \nabla \boldsymbol{u} \|$. The lemmas just proved will be used to obtain estimates for $\| u \|$ and $\| \nabla u \|$ under the assumptions (35) and/or (47), (36) and (37).

In view of (37), the integral in (18) is finite if

$$\lambda + \tfrac{3}{2} < l + 2 \qquad \text{(that is, } \lambda < l + \tfrac{1}{2}\text{)}. \tag{52}$$

Hence, if (16) and (52) hold, then the function in (18) satisfies

$$\chi(r) = O(r^{2\lambda+3}). \tag{53}$$

Then, by (17),

$$\| u - u^N \| = O(r^{\lambda+\frac{3}{2}}) \text{ as } r \rightarrow 0. \tag{54}$$

This estimate and (46) imply that, for every $\epsilon > 0$,

$$\| u \| = o(r^{N-\epsilon}) + O(r^{\lambda+\frac{3}{2}}), \tag{55}$$

when (16), (35)-(37) and (52) hold. If, in addition, (47) holds, then

$$\| u - r^N S_N \| = o(r^N) + O(r^{\lambda+\frac{3}{2}}), \tag{55 bis}$$

for some spherical harmonic $S_N = S_N(e)$ of order N.

If, in the arguments leading to (55), (55 bis), Lemmas 1 and 5 are replaced by Lemmas 3 and 6, one obtains

$$\| u_\phi \| = o(r^{N-\epsilon}) + O(r^{\lambda+\frac{3}{2}}), \tag{56}$$

for every $\epsilon > 0$, when (16), (35)-(37) and (52) hold. If, in addition, (47) holds, this can be improved to

$$\| u_\phi - (r^N S_N)_\phi \| = o(r^N) + O(r^{\lambda+\frac{3}{2}}). \tag{56 bis}$$

Similarly, Lemmas 2 and 7 give

$$r \| u_r \| = o(r^{N-\epsilon}) + O(r^{\lambda+\frac{3}{2}}), \tag{57}$$

for every $\epsilon > 0$, when (16), (35)-(37) and (52) hold. If, in addition, (47) holds, then (57) can be improved to

$$r \| u_r - (r^N S_N)_r \| = o(r^N) + O(r^{\lambda+\frac{3}{2}}). \tag{57 bis}$$

It will now be shown that if (16), (35)-(37) and (52) are satisfied, then, for every $\epsilon > 0$,

$$\| \nabla u \| = O(r^{N-1-\epsilon}) + O(r^{\lambda+\frac{1}{2}}). \tag{58}$$

If (35) is strengthened to (47), this can be improved to

$$\| \nabla u - \nabla (r^N S_N(e)) \| = o(r^{N-1}) + O(r^{\lambda+\frac{1}{2}}). \tag{58 bis}$$

To this end, note that

$$(59_1) \qquad r u_r = u_x x + u_y y + u_z z, \qquad\qquad (59_2) \qquad u_\phi = - u_x y + u_y x,$$

and put

$$u_\psi = u_x z - u_z x. \tag{59$_3$}$$

It is clear that u_ψ satisfies the same type of inequalities, (56) and/or (56 bis), as u_ϕ. If the equations (59) are solved for u_x, one obtains

$$r^2 u_x = x r u_r - y u_\phi + z u_\psi; \text{ hence } r^2 | u_x |^2 \leqq | r u_r |^2 + | u_\phi |^2 + | u_\psi |^2,$$

by Schwarz's inequality. Thus (56), (57) and the analogue of (56) for u_ψ show that $r \| u_x \| = O(r^{N-1-\epsilon}) + O(r^{\lambda+\frac{3}{2}})$ for every $\epsilon > 0$. A similar estimate holds for $\| u_y \|$ and $\| u_z \|$. Thus (58) follows. The relation (58 bis) is obtained in the same way.

7. Proof of Theorem 1. The partial differential equation (1) can be written as (14), where

$$f = - (a \cdot \nabla u + \gamma u). \tag{60}$$

The assumption (4) and the fact that u is of class C^1 imply that, for $N = 1$, there exists a spherical harmonic S_1 (of order 1) satisfying (3). If $S_1(e) \neq 0$, then the proof of Theorem 1 is complete. If $S_1(e) \equiv 0$, then, for $N = 2$,

$$u = o(r^{N-1}), \qquad \nabla u = o(r^{N-2}). \tag{61}$$

It will be shown that if (61) holds for some integer $N(\geqq 2)$, then there exists a spherical harmonic $S_N(e)$ of order N such that (3) holds, where $S_N(e)\equiv 0$ is not excluded.

In view of (60), the assumption (61) implies that (35)-(37), but not (47), are satisfied when $l=N-2$. Thus, if λ is any number satisfying

$$N=l+2>\lambda+\tfrac{3}{2}\,(>\lambda+\tfrac{1}{2}), \tag{62}$$

then the relations (55) and (58) are valid for every $\epsilon>0$. Let $0<\epsilon<1$ and $\lambda+\frac{3}{2}=N-\epsilon$. Then (60), (55) and (58) show that (37) holds if $l=N-1-\epsilon$.

Thus, the conditions for (55 bis) and (58 bis) are satisfied if

$$\lambda+\tfrac{3}{2}<l+2=N+1-\epsilon; \tag{63}$$

cf. (16) and (47). Hence there exists a spherical harmonic S_N of order N satisfying (55 bis). Since $\epsilon>0$ can be chosen arbitrarily small in (63), this means that

$$\| u(r,e)-r^N S_N(e)\| = o(r^N) \text{ as } r\rightarrow 0. \tag{64}$$

Also, (58 bis) gives

$$\| \nabla u(r,e)-\nabla(r^N S_N(e))\| = o(r^{N-1}) \text{ as } r\rightarrow 0. \tag{64 bis}$$

It will be shown that (61) and (64) imply (3). To this end, let $G=G_R(x,\xi)$ denote the Green kernel belonging to Laplace's equation and the boundary condition $u=0$ on the sphere of radius R. Thus,

$$4\pi G=4\pi G_R(x,\xi)=|\xi-x|^{-1}-R\,|\xi^*-x|^{-1}\,|\xi|^{-1}, \tag{65}$$

where $\xi^*=R^2\xi/|\xi|^2$ is the image of ξ under inversion of the sphere $|\xi|=R$. Let $P=P_R(x,Re)$ denote the Poisson kernel belonging to the (surface of the) sphere of radius R; so that

$$P=P_R(x,Re)=(4\pi R)^{-1}(R^2-|x|^2)\,|x-Re|^{-3}. \tag{66}$$

Since $u=u(x)$ is a solution of (14), where $f(x)$ is given by (60), it is seen that, for $r=|x|<R$,

$$u(x)=H_R(x)+v_R(x) \text{ and } \nabla u(x)=\nabla H_R(x)+\nabla v_R(x), \tag{67}$$

where $H_R(x)$ is the harmonic function

$$H_R(x)=\iint_B P(x,Re)u(Re)R^2d\sigma \tag{68}$$

and $v_R(x)$ is the particular solution

$$(69) \qquad v_R(x) = -\iiint\limits_{|\xi|<R} G(x,\xi) f(\xi)\, d\tau$$

of (14). In (69), $d\tau$ is the element of volume in the sphere $|\xi| < R$.

Since $r^N S_N(e)$ is an harmonic function, it follows that

$$(70) \qquad H_R(x) - r^N S_N(x/r) = \iint\limits_{B} P(x, Re)\{u(Re) - R^N S_N(Re)\} R^2 d\sigma.$$

If $|x| \leqq \frac{1}{2}R$, then $|P(x, Re) R^2| \leqq \text{const.}$ Hence, by Schwarz's inequality,

$$|H_R(x) - r^N S_N(x/r)| \leqq \text{Const.}\, \| u(Re) - R^N S_N(e) \| \text{ if } |x| \leqq \tfrac{1}{2}R.$$

Consequently, (64) shows that, as $R \to 0$,

$$(71) \qquad H_R(x) - r^N S_N(x/r) = o(r^N) \text{ uniformly for } |x| \leqq \tfrac{1}{2}R.$$

A similar argument shows that, as $R \to 0$,

$$(72) \qquad \nabla H_R(x) - \nabla (r^N S_N(x/r)) = o(r^{N-1}) \text{ uniformly for } |x| \leqq \tfrac{1}{2}R.$$

The assumption (61) shows that the function $f(x)$ in (60) is $o(r^{N-2})$ as $r \to 0$. Hence, by (69),

$$(73) \qquad |v_R(x)| = o(R^{N-2}) \iiint\limits_{|\xi|<R} |G(x,\xi)|\, d\tau.$$

The inequality $\iiint\limits_{|\xi|<R} |\xi - x|^{-1} d\tau \leqq 8\pi R^2$, where $|x| \leqq R$, and (65) imply that, as $R \to 0$:

$$(74) \qquad v_R(x) = o(R^N) \text{ uniformly for } |x| \leqq \tfrac{1}{2}R.$$

Thus, the first relation in (3) follows from (71), (74) and the first relation in (67).

In order to obtain the asymptotic relation for ∇u in (3), note that

$$(75) \qquad |\nabla v_R(x)| = o(R^{N-2}) \iiint\limits_{|\xi|<R} |\nabla G(x,\xi)|\, d\tau;$$

cf. the remark leading to (73). The inequality $\iiint\limits_{|\xi|<R} |\xi - x|^{-2} d\tau \leqq 8\pi R$ and (65) imply that, as $R \to 0$,

$$(76) \qquad |\nabla v_R(x)| = o(R^{N-1}) \text{ uniformly for } |x| \leqq \tfrac{1}{2}R.$$

Hence, the assertion in (3) concerning ∇u follows from (67) and (72).

This completes the proof of the fact that (61) implies (3) for some $S_N(e)$. Thus, in order to complete the proof of Theorem 1, it remains to be shown that if $u \not\equiv 0$, then it is impossible that

$$(77) \qquad u = o(r^n), \qquad \nabla u = o(r^{n-1}) \qquad \text{for } n = 0, 1, \cdots.$$

In other words, Corollary 1 remains to be verified (and it is clear, from the result just proved, that there is no loss of generality in strengthening the assumption (7) of Corollary 1 to (77)).

8. Proof of Corollary 1. If l and N are any pair of numbers and if (77) holds, then (36) holds and (37) is a consequence of the definition (60) of f. Let λ be the integer satisfying

$$(78) \qquad \lambda + \tfrac{1}{2} < N < \lambda + \tfrac{3}{2}$$

and let $\mu = \lambda$; so that $2N - 2\lambda - 3 \geqq 1$ and $5 - 2(N+1) + 2\mu \geqq 1$. Then the inequalities (17) in Lemma 1 and (44_1) in Lemma 4, with $N - 1$ replaced by N in the latter, are applicable and give

$$(79) \qquad \| u \|^2 \leqq C\{r^{2N} + r^{2\lambda+3} \int_0^1 s^{-2\lambda}\Phi(s)\,ds\},$$

where C is a constant independent of N and λ and depending only on the bounds of u and ∇u.

The inequalities (29), (31) of Lemmas 2, 3 and the inequalities (44_2), (44_3) of Lemma 4, with $N - 1$ replaced by N, imply that

$$(80) \qquad \| \nabla u \|^2 \leqq C\{N^2 r^{2N-2}/(1 - r^2)^3 + r^{2\lambda+1} \int_0^1 s^{-2\lambda}\Phi(s)\,ds\},$$

by the argument involving (59).

The definition of f in (60) and the relations (79), (80) show that, if a constant c is suitably chosen (depending only on the bounds of u, ∇u, a, γ), then

$$(81) \qquad \Phi(r) \leqq c\{N^2 r^{2N-2}/(1 - r^2)^3 + r^{2\lambda+1} \int_0^1 s^{-2\lambda}\Phi(s)\,ds\}.$$

If this inequality is multiplied by $r^{-2\lambda}$, a quadrature gives

$$\int_0^r s^{-2\lambda}\Phi(s)\,ds \leqq c\{N^2(2N - 2\lambda - 1)^{-1} r^{2N-2\lambda-1}/\epsilon^3 + \tfrac{1}{2} r^2 \int_0^1 e^{-2\lambda}\Phi(s)\,ds\},$$

if $0 \leqq r \leqq (1-\epsilon)^{\frac{1}{2}}$. Thus, $2N-2\lambda-1>2$ implies that

$$(82) \qquad (1-\tfrac{1}{2}cr^2)\int_0^r s^{-2\lambda}\Phi(s)\,ds \leqq c\{N^2r^2/2\epsilon^3+\tfrac{1}{2}r^2\int_r^1 s^{-2\lambda}\Phi(s)\,ds\},$$

if $0 \leqq r \leqq (1-\epsilon)^{\frac{1}{2}}$.

Note that if $f(x)\equiv 0$ in (14), (60), then $u(x)\equiv 0$, for otherwise u is harmonic and satisfies (77). Hence $u(x)\equiv 0$ implies $\Phi(r)\equiv 0$ and conversely.

It will be shown that (77) implies that $u(x)\equiv 0$ on every sphere $|x|\leqq r$ with a radius r satisfying $1-\frac{1}{2}cr^2>0$. Since c is constant, independent of N, λ and depending only on the bounds of u, ∇u, α, γ on $|x|\leqq 1$, it will follow that $u(x)\equiv 0$ on $|x|\leqq 1$.

Suppose, if possible, that $u(x)\not\equiv 0$ on $|x|\leqq r$. Then $\Phi(r)\neq 0$ for some $r=r_0$, where $0\leqq r_0<r$. Let $\epsilon>0$ be chosen so that $r\leqq(1-\epsilon)^{\frac{1}{2}}$. Then (82) is applicable. Since $\Phi(r_0)>0$, it follows that the left side of (82) is minorized by const. $r_0^{-2\lambda}$, where const. >0. But the right side is majorized by Const. $(N^2r^2+r^{-2\lambda})$. In view of (78) and $r_0<r$, this leads to a contradiction if $\lambda\rightarrow\infty$.

This completes the proof of Corollary 1 and of Theorem 1 (hence of Corollary 2 also).

Part 2. Solutions having a singularity.

9. The nature of the singularity. This part of the paper will deal with functions $u=u(x)$ which are C^1-solutions of (1) on a punctured vicinity of $x=0$, say on

$$(83) \qquad 0<|x|\leqq 1.$$

The theorems to follow will consider the three cases in which the solution $u=u(x)$ satisfies one of the following conditions, as $r=|x|\rightarrow 0$:

$$(84) \qquad u(x)=o(1/r),$$

$$(85) \qquad u(x)\rightarrow\infty,$$

$$(86) \qquad u(x)=o(r^{-N-2}), \qquad \nabla u(x)=o(r^{-N-3}) \qquad \text{for some integer } N\geqq 0.$$

The theorem concerning (84) is a theorem on "removable singularities":

THEOREM 2. *Let x, $\gamma(x)$, $\alpha(x)$ satisfy the conditions of Theorem* 1.

Let $u(x)$ be a C^1-solution of (1) *on* (83) *satisfying* (84) *as $r \to 0$. Then $u(0)$ can be defined in such a way that $u(x)$ becomes of class C^1 on $|x| \leqq 1$.*

It will be clear from the proof that the analogue of this theorem is true in the plane. This two-dimensional analogue is an improvement of what follows from the results of [4] on removable singularities, inasmuch as there is no condition on ∇u as $x \to 0$ in Theorem 2. Following Bôcher [1], the theorem on removable singularities is obtained as a consequence of theorems on the existence and uniqueness of a solution having a Green singularity at a given point.

If (84) is replaced by (85), Theorem 2 is altered to

THEOREM 3. *Let x, $\gamma(x)$ satisfy the conditions of Theorem* 1. *Let $u(x)$ be a C^1-solution of* (1) *on* (83) *satisfying* (85), *as $r \to 0$. Then there exists a constant $c > 0$ such that, as $r \to 0$,*

$$(87) \qquad u(x) = c/r + O(|\log r|), \qquad \nabla u(x) = \nabla(c/r) + O(1/r).$$

Bôcher [1] has proved theorems similar to Theorems 2 and 3. In his analogue of Theorem 3, the assertion (87) is considerably improved. Bôcher assumes that the coefficient functions α, γ are analytic, but his proof does not use the full force of this assumption. It requires merely that (1) has an adjoint and that both (1) and its adjoint possess Green functions (for some sphere $|x| \leqq R$).

The proof of Theorem 3 will be similar to the proof of its analogue in the plane; cf. [5], pp. 351-352. Note that the assertion of the two-dimensional theorem (*) in [5] is to the effect that $\nabla u = -c\nabla(\log r) + o(1/r)$, but the proof given there shows that this can be improved to

$$\nabla u = -c\nabla(\log r) + O(|\log r|),$$

cf. [5], p. 358.

Theorem 3 is a uniqueness theorem on Green's singularity. There is an existence theorem analogous to the (two-dimensional) theorem (**) in [5], pp. 360-361.

Theorem 2 will depend on Theorem 3, the existence theorem just mentioned and the methods of Part 1 above.

THEOREM 4. *Let x, $\gamma(x)$, $\alpha(x)$ satisfy the conditions of Theorem* 1. *Let $u(x)$ be a C^1-solution of* (1) *on* (83) *satisfying* (86), *as $r \to 0$. Then there exists a spherical harmonic $S_N(e)$ of order N (possibly $S_N(e) \equiv 0$) satisfying, as $r \to 0$,*

$$u(x) = r^{-N-1}S_N(e) + o(r^{-N-1}), \quad \nabla u(x) = \nabla(r^{-N-1}S_N(e)) + o(r^{-N-2}),$$

uniformly in e.

The proofs of Theorems 2, 3 and 4 will be applicable if the linear equation (1) is replaced by a non-linear equation (5) for which (6) holds. Although the notation of vector products is used in the proof of Theorem 3, it will be clear that this does not affect the validity of the method if the dimension number 3 is increased.

10. Proof of Theorem 4. The proof is similar to that of Theorem 1 and will only be indicated. Since $U = U_{nj}$ need not be continuous at the origin, (20) must be replaced by

$$(88) \qquad (2n+1)U = r^n\Big(A - \int_r^1 s^{1-n}F(s)\,ds\Big) - r^{-1-n}\Big(B - \int_r^1 s^{n+2}F(s)\,ds\Big),$$

where $A = A_{nj}$, $B = B_{nj}$ are integration constants to be determined. Note that the first part of (86) implies that

$$(89) \qquad U = o(r^{-N-2}) \text{ as } r \to 0.$$

The assumptions (86) imply that f in (14), (60) satisfies

$$(90) \qquad f = o(r^{-N-3}) \text{ as } r \to 0,$$

hence

$$(91) \qquad \Phi = o(r^{-2(N+3)}) \text{ and } F = F_{nj} = o(r^{-N-3}).$$

This estimate of F shows that

$$(92) \qquad r^n \int_r^1 s^{1-n}F(s)\,ds = o(r^{-N-1}) \text{ for } n = 0, 1, \cdots.$$

In fact, the expression on the left is $o(|\log r|)$ or $o(r^{-N-1})$ according as (N, n) is or is not $(-1, 0)$.

In view of (89) and (92), the coefficient of r^{-1-n} in (88) is $o(1)$, as $r \to 0$, if $n > N$; so that

(93) (88) reduces to (20) if $n > N$.

The (absolute) convergence of the integral which is the factor of r^{-1-n} in (20) is assured by the last part of (91).

It follows that Lemma 1 remains valid if the conditions $\lambda \geqq 0$ and (16) are replaced by

$$(94) \qquad -\tfrac{3}{2} > \lambda > -N - \tfrac{5}{2}$$

and (18) is replaced by

$$(95) \qquad \chi(r) = r^{2N+2} + r^{2\lambda+3}(2N+2\lambda+5)^{-1}\int_0^1 s^{-2\lambda}\Phi(s)\,ds.$$

Corresponding analogues of Lemmas 2 and 3 hold.

In order to examine the analogues of the lemmas of Section 5, assume that

$$(96) \qquad \Phi(r) = o(r^{-2(l+1)}), \text{ where } l \leqq N+2.$$

This corresponds to (37); the analogue, $\| u(r,e) \| = o(r^{-N-2})$, of (35) is a consequence of (86). It follows that

$$r^{-1-n}\int_r^1 s^{n+2}F(s)\,ds \text{ is } o(r^{1-l}) \text{ or } o(r^{-1-n}\,|\log r\,|)$$

according as $n < l - 2$ or $n = l - 2$. Also

$$\lim_{r\to 0}\,(B - \int_r^1 s^{n+2}F(s)\,ds) \text{ exists if } n > l-2.$$

Thus, (96) implies that, as $r \to 0$,

$$\| u^N \| = o(r^{-N-1-\epsilon}) \text{ if } N \geqq 0 \text{ and } \epsilon > 0,$$

and this can be improved to

$$\| u^N - r^{-1-N}S_N(e) \| = o(r^{-1-N}) \text{ if } l < N+2,$$

where S_N is some spherical harmonic of order N. These assertions correspond to Lemma 5. It is clear that assertions analogous to Lemmas 6 and 7 hold.

Thus, the arguments in the first part of Section 7 imply the existence of a spherical harmonic S_N of order N satisfying, as $r \to 0$,

$$(97) \quad \| u(x) - r^{-1-N}S_N(e) \| = o(r^{-1-N}), \quad \| \nabla u(x) - \nabla(r^{-1-N}S_N(e)) \| = o(r^{-2-N})$$

if $N \geqq 0$. The proof of Theorem 4 can now be concluded by a modification of the arguments of the second part of Section 7. The Green kernel (65) of the sphere $|x| \leqq R$ must be replaced by the Green kernel of a region of the type $R/4 \leqq |x| \leqq 4R$, and the Poisson kernel (66) by the kernel used to determine an harmonic function on $R/4 < |x| < 4R$ when its values on the two concentric spheres $|x| = R/4$ and $|x| = 4R$ are given; cf., e.g., [3]. The analogues of (71), (72), (74), (76), when $|x| \leqq \frac{1}{2}R$ is replaced by $\frac{1}{2}R \leqq |x| \leqq 2R$, can then be deduced.

11. Proof of Theorem 3. The proof will be similar to the proof for the analogous theorem in the plane; cf. [5]. Assume first that

$$\gamma(x) \equiv 0 \tag{98}$$

in (1). Then, by Theorem 1, the zeros of the gradient of u can cluster only at $x = 0$, since (98) shows that Corollary 2 of Theorem 1 is applicable to $u(x) - u(x_0)$ at any point $x = x_0$ of (83). It follows, by arguing as in [5], p. 355, that there exist arbitrarily large constants C such that the locus

$$L = L_C \colon\ u(x) = C \tag{99}$$

is a closed (Jordan) C^1-surface surrounding $x = 0$. Furthermore,

$$u_p = p \cdot \nabla u \geqq 0 \text{ on } L, \tag{100}$$

where p is the inward unit normal vector to L (and u_p is, therefore, the corresponding normal derivative of u).

If T is a suitable region bounded by a surface S and p is the unit outward normal on S, finally v is a suitable scalar function, then Green's identity is

$$\iint_S v \nabla u \cdot p\, dS = \iiint_T (v \Delta u + \nabla v \cdot \nabla u)\, dT. \tag{101}$$

Also, if w is a smooth vector function, Green's identity gives

$$\iint_S (w \times \nabla u) \cdot p\, dS = \iiint_T \nabla u \cdot (\nabla \times w)\, dT,$$

where $\times$ denotes vector multiplication. If the second of these relations is subtracted from the first, one obtains

$$\iint_S \nabla u \cdot (vp + (w \times p))\, dS = \iiint_T v \Delta u\, dT, \tag{102}$$

provided that

$$\nabla \times w = \nabla v.$$

For a given v, the last equation has a solution w if and only if v is harmonic. If u is a C^1-solution of a Poisson equation (14) and if S and w are sufficiently smooth, then (102) remains valid.

Let ξ be a 3-dimensional vector and let $x - \xi$ have the components (X, Y, Z). Choose

$$v = X/\rho^3, \text{ where } \rho = |x - \xi| \tag{103}$$

(so that v is the first component of $-\nabla(1/\rho)$). A vector w satisfying the equation following (102) is

$$w = (0,\ Z/\rho^3,\ -Y/\rho^3). \tag{104}$$

Apply the relation (102) to the region T bounded by the sphere B: $|\xi| = 1$, the locus (99) and a small sphere $|x - \xi| = \epsilon$ about a fixed point x (between B and L). On the sphere $|x - \xi| = \epsilon$, the outward unit normal p is $(x - \xi)/\epsilon$. Hence, $vp + (w \times p) = (1/\epsilon^2, 0, 0)$, and so

$$\lim_{\epsilon \to 0} \iint\limits_{|x-\xi|=\epsilon} \nabla u \cdot (vp + (w \times p))\, dS = 4\pi u_x(x)$$

(here $u_x = \partial u/\partial x$ is the first component of ∇u). Also, note that $\nabla u \cdot (w \times p) = 0$ on L. Hence (102)-(104) lead to

$$4\pi u_x(x) = -\iint\limits_{L} v u_p\, dS + \iint\limits_{B} \nabla u \cdot (vp + (w \times p))\, d\sigma - \iiint\limits_{T} v \Delta u\, d\tau. \tag{105_x}$$

where $T = T(C)$ is the region bounded by B and $L = L_C$.

The relations (105_x), the corresponding relations (105_y), (105_z) for $u_y(x)$, $u_y(x)$ and Green's identity

$$\iint\limits_{L} u_p\, dS = -\iint\limits_{B} u_p\, d\sigma + \iiint\limits_{T} \Delta u\, d\tau \tag{106}$$

are the exact analogues of formulae (22), (28) used in [5] for the proof of the planar case of Theorem 3 (when (98) is assumed). A modification of the arguments of [5], pp. 356-357, leads to the result that

$$4\pi c = \lim_{C \to \infty} \iint\limits_{L} u_p\, dS \text{ exists} \tag{107}$$

and to the asymptotic relation

$$\nabla u = c\nabla(1/r) + O(1/r) \text{ as } r \to 0. \tag{108}$$

The relation (108) and a quadrature of (108) give Theorem 3, in the case (98). Finally, the general case can be reduced to (98) by a variation of constants; cf. [5], pp. 359-360.

12. Proof of Theorem 2. The existence theorem mentioned after Theorem 3 is to the effect that if $R > 0$ is sufficiently small, then (1) has a C^1-solution $u = G(x)$ on $0 < |x| \leqq R$ satisfying $G(x) \to \infty$ as $x \to 0$. By Theorem 3,

$$G(x) = c/r + O(|\log r|), \qquad \nabla G(x) = \nabla(c/r) + O(1/r),$$

as $r \to 0$, where $c > 0$.

If (84) holds, then $u(x) + G(x)$ is a C^1-solution of (1) on $0 < |x| \leqq R$ satisfying $u(x) + G(x) \to \infty$ as $x \to 0$. It follows from Theorem 3, (84) and the last formula line that

$$(109) \qquad u(x) = O(|\log r|), \qquad \nabla u(x) = O(1/r),$$

as $r \to 0$. Thus

$$(110) \quad U = U_{nj} = O(|\log r|), \quad (111) \quad \Phi = O(r^{-2}) \text{ and } F = F_{nj} = O(r^{-1}).$$

One of the terms on the right of (88) is

$$(112) \qquad -r^n \int_r^1 s^{1-n} F(s)\,ds = O(1), \ O(r\,|\log r|) \text{ or } O(r)$$

according as $n = 0$, $n = 1$ or $n > 1$. In view of (110), the coefficient of r^{-1-n} in (88) is $o(1)$ as $r \to 0$. Hence

$$B - \int_r^1 s^{n+2} F(s)\,ds = \int_0^r s^{n+2} F(s)\,ds,$$

where the absolute convergence of the last integral is assured by (111). Thus

$$(113) \qquad (88) \text{ reduces to } (20) \text{ if } n \geqq 0.$$

Analogues of the Lemmas 1-3 give $\| u - u^0 \|^2 = O(r^2 + r^{2\lambda+3})$ if $\lambda < -\frac{1}{2}$. The case $n = 0$ of (20) and the estimate for F in (111) show that $U_{00} = \text{Const.} + O(r)$. Hence $\| u - \text{Const.} \| = O(r)$ as $r \to 0$.

The proof of Theorem 2 can now be completed along the lines used to prove Theorem 1. In fact, the proof of

$$u(r, e) - \text{Const.} = O(r)$$

is identical with the proof of the first formula of (3) in Theorem 1; cf. the derivation of (71), (74). The analogue of (72), namely, $\nabla H_R(x) = O(1)$ uniformly for $|x| \leqq \frac{1}{2}R$, is clearly valid. In order to obtain the analogue of (76), namely, $\nabla v_R(x) = O(1)$ uniformly for $|x| \leqq \frac{1}{2}R$, replace (75) by $|\nabla v_R(x)| \leqq I_1 + I_2$, where I_1, I_2 are the integrals of $|\nabla G(x, \xi)|\,|\nabla u|$ over the domains $|\xi| \leqq R/4$ and $R/4 \leqq |\xi| \leqq R$, and let $|x| = \frac{1}{2}R$. Then, since $|\nabla u| = O(1/r)$, by (1) and (109),

$$I_1 = O(R^{-2}) \int\int\int |\xi|^{-1} d\tau = O(1),$$

$$I_2 = O(R^{-1}) \int\int\int |x - \xi|^{-2} d\tau = O(1).$$

Consequently, $|\nabla u(x)| = O(1)$ as $|x| = \frac{1}{2}R \rightarrow 0$. Thus $f(x)$ in (14), (60) is bounded. From this, it follows that $u(x)$ is of class C^1 (at $x = 0$ also). This proves Theorem 2.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] M. Bôcher, "Singular points of functions which satisfy partial differential equations of the elliptic type," *Bulletin of the American Mathematical Society*, vol. 9 (1903), pp. 455-465.

[2] T. Carleman, "Sur les systèmes linéaires aux dérivées partielles du premier ordre à deux variables," *Comptes Rendus*, vol. 197 (1933), pp. 471-474.

[3] U. Dini, "Il problema di Dirichlet in un'area anulare, e nello spazio compreso fra due sfere concentriche," *Rendiconti del Circolo Matematico di Palermo*, vol. 16 (1913), pp. 1-28.

[4] P. Hartman and A. Wintner, "On the local behavior of solutions of non-parabolic partial differential equations," *American Journal of Mathematics*, vol. 75 (1953), pp. 449-476.

[5] ——— and A. Wintner, "On the local behavior of solutions of non-parabolic partial differential equations. II. The uniqueness of the Green singularity," *ibid.*, vol. 76 (1954), pp. 351-361.

[6] C. Müller, "On the behavior of the solutions of the differential equations $\Delta U = F(x, U)$ in the neighborhood of a point," *Communications on Pure and Applied Mathematics*, vol. 7 (1954), pp. 505-516.

ON THE ASSIGNMENT OF ASYMPTOTIC VALUES FOR THE SOLUTIONS OF LINEAR DIFFERENTIAL EQUATIONS OF SECOND ORDER.*

By PHILIP HARTMAN and AUREL WINTNER.

Let both coefficients of the homogeneous, linear differential equations

$$x'' + g(t)x' + f(t)x = 0 \tag{1}$$

be given as continuous functions for large positive t.

Suppose that, corresponding to every given value, c, the differential equation (1) has a unique solution $x = x(t)$ satisfying

$$x \rightarrow c \tag{2}$$

as $t \rightarrow \infty$ (so that $x(\infty)$ can be assigned as an "initial condition" and this single integration constant determines the solution $x(t)$ uniquely, even though (1) is of second order). Then (1) will be called *of type* (*). Clearly, (1) is of type (*) if and only if (1) has exactly one solution satisfying the case $c = 1$ of (2) and, in addition, the case $c = 0$ of (2) is satisfied only by the trivial solution ($\equiv 0$). It is also clear that, since the general solution of (1) contains two arbitrary constants, $x(\infty)$ cannot exist (as a finite limit) for all solutions $x(t)$, if (1) is of type (*).

If $x(\infty)$ exists (as a finite limit) for all solutions and if $x(\infty) \neq 0$ holds for some solution, then (1) will be called *of type* (**). Clearly, this will be the case if and only if there exist a solution satisfying the case $c = 1$ of (2) and a non-trivial solution ($\not\equiv 0$) satisfying the case $c = 0$ of (2). By adding to the former any constant multiple of the latter, it is seen that (1) is of type (**) if and only if there belongs to some and/or every $c \neq 0$ two, linearly independent, solutions $x(t)$ satisfying (2).

The following theorem will first be proved:

(I) *In order that* (1) *be of type* (*), *it is sufficient that*

(i) *the coefficient functions of* (1) *satisfy the conditions*

$$(3_1) \qquad \int^{\infty} t\,|f(t)|\,dt < \infty, \qquad\qquad (3_2) \qquad g(t) \leqq 0;$$

and it is also sufficient that

---

* Received January 3, 1955.

(ii) *condition* (3_1) *of* (i) *be reduced, but condition* (3_2) *of* (i) *be strengthened, as follows:*

$$(4_1)\qquad \int^{\infty} |f(t)|\,dt < \infty, \qquad\qquad (4_2)\qquad g(t) \leqq \text{const.} < 0.$$

Condition (3_1) of (i) cannot be relaxed to (4_1) if (4_2) is relaxed to (3_2), as in (i). Otherwise it would follow that (1) is of type (*) whenever (4_1) holds and $g(t) \equiv 0$. But this is not true. In fact, it is known that if $f(t)$ is real-valued and does not change sign, then (3_1) is not only sufficient (Bôcher) but necessary (Weyl) as well in order that the case $g(t) \equiv 0$ of (1) be of type (*) (for a simple proof, cf. [6]).

It will be clear from the proof of part (i) of (I) that the assumption (3_2) of (i) can be generalized to

$$(5)\qquad \int^{\infty} g^{+}(t)\,dt < \infty, \quad \text{where } g^{+} = \max(0, g);$$

so that (3_1) and (5) (hence either (3_1) and $\int^{\infty} |g(t)|\,dt < \infty$ or (3_1) and (3_2)) are sufficient in order that (1) be of type (*). Note that g (but not f) is here assumed to be real-valued; cf. the remark preceding the proof of (I) below.

The proof will show that, under the conditions of (I), it is possible to assert much more than the fact that (1) is of type (*). As will be seen from the proof of (I), condition (3) or (4) implies that (1) has a pair of linearly independent solutions $x = x_1(t)$, $x_2(t)$ satisfying, as $t \to \infty$,

$$x_1(t) \to 1, \quad x_1' = o(1/s); \qquad x_1(t) \sim s, \quad x_2' \sim \mu,$$

where $\mu = \mu(t)$ and $s = s(t)$ are defined by (10) and (11) below.

A simple illustration of part (ii) of (I) is the following known fact: Whenever (4_1) is satisfied,

$$(6)\qquad z'' - \{1 + f(t)\}z = 0$$

has a (unique) solution $z(t)$ which is asymptotically equal to e^{-t}, that is, for which the function $x(t) = e^{t}z(t)$ satisfies the case $c = 1$ of (2). In order to see this, it is sufficient to observe that substitution of $z = e^{-t}x$ into (6) leads to the case $g(t) \equiv -2$ of (1); so that part (ii) of (I) becomes applicable.

A criterion which belongs to property (**) in the same way as part (ii) of (I) belongs to property (*) is as follows:

(II) *In order that* (1) *be of type* (**), *it is sufficient that the coefficient functions of* (1) *satisfy the conditions*

$$(7_1) \qquad \int^{\infty} |f(t)dt < \infty, \qquad\qquad (7_2) \qquad \operatorname{Re} g(t) \geqq \text{const.} > 0.$$

Note that, while (7_1) is identical with (4_1), condition (7_2) is a dual of condition (4_2), if $g(t)$ is real-valued (which is not assumed in (II)). Since (II) corresponds to part (ii) of (I), it is worth mentioning that what would correspond to part (i) of (I) as a (**)-criterion is false. In other words, (3_1) and $g(t) \geqq 0$ together fail to assure that (1) is of type (**). This is seen by choosing $f(t) \equiv 0$ and $g(t) \equiv 0$.

It should also be remarked that (I) becomes false if $g(t)$ is allowed to be complex-valued and $g(t)$ is replaced by $\operatorname{Re} g(t)$ in (3_2) and/or (4_2). In order to see this, consider the case $f(t) \equiv 0$ of (1) (so that f satisfies (3_1)). The general solution of the corresponding differential equation (1) is $c_1 + c_2 s(t)$, where c_1, c_2 are arbitrary constants and $s(t)$ is defined by (10), (11) below. Thus the desired example results if $\operatorname{Re} g(t) = -1$ and $s(\infty) = \lim s(t)$ exists as a finite limit (for, in this case, (1) is of type (**), not (*)). If $g(t) = -(1 + ite^t)$, then

$$s(t) = \int^{t} \mu(u)\, du \quad \text{and} \quad \mu(u) = e^u \exp i \int^{u} ve^v\, dv.$$

Define $w = w(u) = \int^{u} ve^v\, dv$, so that $e^u\, du = dw/u$. The change of integration variables $u \to w$ gives for $s(t)$ the formula

$$s(t) = \int^{w(t)} (e^{iw}/u(w))\, dw,$$

where $u = u(w)$ is the inverse of $w = w(u)$. It is clear that $w(t) \to \infty$ as $t \to \infty$ and that $1/u(w)$ tends monotonously to 0 as $w \to \infty$. Hence $s(\infty)$ exists (as a finite limit). Thus the case $f \equiv 0$, $g = -(1 + ite^t)$ of (1) is not of class (*).

Proof of (I). It is known that, if $h(s)$ is defined and continuous for large positive s, then the condition

$$(8) \qquad\qquad \int^{\infty} s\,|\,h(s)\,|\, ds < \infty$$

(and even a somewhat weaker assumption; cf. [5], [3]) is sufficient in order that the differential equation

$$\ddot{x} + h(s)x = 0, \tag{9}$$

where $\dot{x} = dx/ds$, has two solutions, say $x = x_1(s)$ and $x = x_2(s)$, satisfying $x_1(s) \to 1$, $x_2(s) \sim s$ (and $\dot{x}_1 = o(1/s)$, $\dot{x}_2 \to 1$) as $s \to \infty$. Hence (8) is sufficient in order that (9) be of type (*). Both assertions of (I) will be reduced to this fact.

First, whether (4_2) or just (3_2) is assumed, it is clear that $\mu = \mu(t)$, where

$$\mu = \exp - \int^t g(u)\,du, \tag{10}$$

is a positive function, hence $s = s(t)$, where

$$ds = \mu dt, \tag{11}$$

is an increasing function, and that $s \to \infty$ as $t \to \infty$ (the choice of the integration constants in (10) and (11) will be immaterial). On the other hand, if the independent variable t is replaced by s in (1), it is readily verified from (10) and (11) that (1), where $x' = dx/dt$, goes over into (9), where $\dot{x} = dx/ds$ and

$$h = f/\mu^2. \tag{12}$$

Since (8) implies that (9) is of type (*), it follows that (1) will be proved to be of type (*) if it is shown that, whether (3_1) and (3_2) or (4_1) and (4_2) are assumed, the function (12) of s will satisfy (8). Since $s = \infty$ corresponds to $t = \infty$, this requires that

$$\int^\infty s(t)\,|f(t)|/\mu^2(t)\,ds(t) < \infty$$

or, what in view of (11) and (10) is the same thing, that

$$\int^\infty |f(t)| \left(\int^t \mu(v)\,dv\right)/\mu(t)\,dt < \infty. \tag{13}$$

Ad (i). It is clear from (3_2) and (10) that $\mu(t)$ increases with t. This implies that the indefinite integral of $\mu(t)$ is majorized by a constant multiple of $t\mu(t)$. Hence (13) follows from (3_1).

Ad (ii). If only (4_1) is assumed, (13) will hold if the quotient

$$\left(\int^t \mu(v)\,dv\right)/\mu(t) \tag{14}$$

is bounded as $t \to \infty$. On the other hand, in view of an applicable "O-variant" of l'Hôpital's rule, (14) will be bounded if the quotient $\mu(t)/\mu'(t)$ is. But (10) shows that the latter quotient is identical with the function $-1/g(t)$. Finally, the boundedness of this function is assured by (4_2).

Proof of (II). It is known ([4], pp. 55-56) that if $p(t) \neq 0$ and $q(t)$ are, for large positive t, continuous functions satisfying the pair of conditions

$$(15_1) \quad \int^{\infty} |p(t)|^{-1}\,dt < \infty, \qquad (15_2) \quad \int^{\infty} |q(t)|\Big(\int_t^{\infty} |p(u)|^{-1}\,du\Big)dt < \infty,$$

then the differential equation

$$(16) \qquad (p(t)x')' + q(t)x = 0$$

has a pair of linearly independent solutions, say $x = x_1(t)$ and $x = x_2(t)$, satisfying $x_1(t) \to 1$ and $x_2(t) \to 0$ as $t \to \infty$. This means that (15_1) and (15_2), where $p(t) \neq 0$, imply that (16) is of type (**).

On the other hand, if (1) is multiplied by the (non-vanishing) function $e^{j(t)}$, where

$$(17) \qquad j(t) = \int^{t} g(u)\,du,$$

then (1) appears in the form (16), where

$$(18_1) \quad p(t) = e^{j(t)}, \qquad (18_2) \quad q(t) = f(t)e^{j(t)}.$$

Hence (II) will be proved if it is shown that (7_1) and (7_2) imply (15_1) and (15_2) by virtue of (17)-(18_2); in other words, that

$$(19_1) \qquad \int^{\infty} |e^{-j(t)}|\,dt < \infty,$$

$$(19_2) \qquad \int^{\infty} |f(t)e^{j(t)}|\Big(\int_t^{\infty} |e^{-j(u)}|\,du\Big)dt < \infty.$$

If $c > 0$ denotes the constant occurring in (7_2), it is clear from (17) that $|e^{-j(t)}|$ is majorized by a constant multiple of e^{-ct}, which implies that (19_1) is satisfied. But it is also seen from (7_2) and (17) that

$$\int_t^{\infty} |e^{-j(u)}|\,du \leq c^{-1} \int_t^{\infty} |\operatorname{Re} g(u)|\,|e^{-j(u)}|\,du \leq c^{-1}\,|e^{-j(t)}|.$$

Hence (19_2) is satisfied if $\int^{\infty} |f(t)e^{j(t)}|\, c^{-1}\, |e^{-j(t)}|\, dt < \infty$ is, that is, if (7_1) is.

A theorem of somewhat different nature will now be considered.

(III) *In the differential equation*

$$z'' + f(t)z = 0, \tag{20}$$

let $f(t)$, *where* $0 \leqq t < \infty$, *be a complex-valued, continuous function having a non-positive real part. Then* (20) *has a solution* $z(t) \not\equiv 0$ *corresponding to which* $|z(t)|^2$ *is a non-increasing, convex function on* $0 \leqq t < \infty$. *In particular, the limit of* $|z(t)|$ *as* $t \rightarrow \infty$ *exists as a finite limit* $(\geqq 0)$. *Any solution* $z = w(t)$ *of* (20) *linearly independent of this solution* $z(t)$ *is subject to the condition* $|w(t)| \rightarrow \infty$ *as* $t \rightarrow \infty$.

If $f(t)$ is real-valued, then (III) reduces to a well known theorem of A. Kneser. In order to prove (III), let $z = x + iy$ and $f(t) = a(t) + ib(t)$. Then (20) can be written as a real system,

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} + F(t)\begin{pmatrix} x \\ y \end{pmatrix} = 0, \text{ where } F = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Since $a \leqq 0$ by assumption, the symmetric part of the matrix F is non-positive definite. Hence (III) can be concluded from a result of [5].

Appendix.

When confined to the real field, both (I) and (II) refer to cases in which (1) is non-oscillatory (cf. below). This Appendix deals with the general class of non-oscillatory differential equations (1) under the assumption that the x'-term is absent from (1).

For large positive t. let $f(t)$ be a real-valued, continuous function, and consider only those solutions $x(t)$ of the differential equation

$$x'' + f(t)x = 0 \tag{1}$$

which are real-valued and distinct from the trivial solution $(\not\equiv 0)$. Thus, since the values of x and x' at a given t determine a solution of (1) uniquely, x and x' cannot vanish simultaneously, and so $x(t)$ must change sign wherever it vanishes. If there exists a particular solution $x(t)$ which does not vanish for large t, say for $t_0 \leqq t < \infty$, then (1) is called non-oscillatory, and otherwise oscillatory. This classification of the differential equations (1) is independent of the choice of the particular solution $x(t)$, since. in view of Sturm's separation theorem, every solution of (1) must have zeros clustering at $t=\infty$ if one solution has such zeros.

If (1) is non-oscillatory, then, for reasons which will be apparent from the results to be proved below, a solution $x(t)$ of (1) will be called *principal* or *non-principal* according as the (continuous) function $1/x(t)$ is not or is of class L^2 on the half-line $[t_0, \infty)$ when t_0 is large enough (namely, larger than the last zero of $x(t)$).

If $x(t)$ is a non-principal solution, that is, if the second factor of the product

$$(2) \qquad x(t) \int_t^\infty (x(s))^{-2}\,ds$$

can be formed (for $t \geqq t_0$). then two differeniations of the product show that the function (2) is a solution of (1), since $x(t)$ is. These two solutions are linearly independent. since their ratio is not a constant. Actually, the solution (2) can be obtained by solving for $y=y(t)$ the case $c=-1$ of the differential relation (3) below.

Let $x=x(t)$ and $x=y(t)$ be any two solutions of (1). Then their Wronskian is a constant,

$$(3) \qquad xy'-yx'=c,$$

and the constant c is distinct from 0 if and only if the two solutions are linearly independent. Suppose that they are, that (1) is non-oscillatory, and that t is large enough. Then (3) can be written, on the one hand. as

$$(4) \qquad (\log y)'-(\log x)'=c/(xy), \quad \text{where } c\neq 0,$$

and, on the other hand, as

$$(5) \qquad (y/x)'=c/x^2, \quad \text{where } c\neq 0.$$

With the aid of these formulations of (3) (and of its formulation (2), specified above), the following theorem will be proved:

THEOREM. *If* (1) *is non-oscillatory, then it has a principal solution. and the latter is unique to an arbitrary constant factor* ($\neq 0$).

This contains the following corollary (which corollary of the theorem will, however, be needed in the proof of the theorem):

COROLLARY. *If* (1) *is non-oscillatory, then one* (*at least*) *of any two of its linearly independent solutions is non-principal.*

This, in turn, implies the existence of a non-principal solution. A proof for the existence of such a solution is between the lines of [1], p. 703 (and the properties of non-principal solutions are further analyzed in [2], p. 633). The following proof of the Corollary is a simplification of the proof just mentioned.

Let $x(t)$ and $y(t)$ be the two solutions referred to in the Corollary. Then, since (5) holds for large t, the function $y(t)/x(t)$ is ultimately monotone. Hence

$$y(t)/x(t) \to C, \quad \text{as } t \to \infty, \tag{6}$$

holds for some limit C, provided that $C = \infty$ and $C = -\infty$ are allowed. The latter two possibilities can, however, be reduced to the case $C = 0$, since $x(t)$ and $y(t)$ can be interchanged. But if (6) holds for some finite C, then it is seen from (5) that an indefinite integral of the square of $1/x(t)$ must tend to a finite limit as $t \to \infty$. Since this means that $x(t)$ is non-principal, the assertion of the Corollary follows.

Besides the Corollary, part of the following Lemma will be needed in the proof of the Theorem.

Lemma. *If* (1) *is non-oscillatory and if* $x(t)$ *is a fixed non-principal solution of* (1), *then every solution of* (1) *is of the form*

$$Cx(t) + o(|x(t)|), \qquad \text{as } t \to \infty, \tag{7}$$

where the constant C *is or is not* 0 *according as the solution* (7) *is principal or non-principal.*

In order to prove this, suppose that $y(t)$ is a non-principal solution. Then, since $x(t)$ is supposed to be non-principal, and since the product of two functions of class L^2 is of class L, it follows from (4) that $\log y - \log x$ tends to a finite limit as $t \to \infty$, which means that (6) holds for a finite $C \neq 0$. This proves that assertion of the Lemma which concerns the case $C \neq 0$ of (7). In the remaining case, the assertion of (7) is that every principal solution must be $o(|x(t)|)$ if $x(t)$ is non-principal.

If $x(t)$ is non-principal, then the solution (2) exists (for large t) and is of the form $x(t)o(1)$, which is $o(|x(t)|)$. Hence, by that part of the Lemma which has already been proved, the solution (2) cannot be non-principal. Consequently, the proof of the Lemma as well as the proof of the Theorem will be complete if it is assured that there cannot exist two linearly independent solutions both of which are principal. But this is assured by the Corollary.

REMARK. (*Added* 4. 8. 1955). The notion of a principal solution has a curious application to the general case of (1), the case in which $f(t)$ is real-valued and continuous for large positive t but (1) is not required to be non-oscillatory.

First, it is easily verified from (3), (1) and the case $x=y$ of (1) that $r=(x^2+y^2)^{\frac{1}{2}}>0$ is a solution of

$$r''+g(t)r=0, \tag{8}$$

where $g(t)=f(t)-c^2/r^4(t)$, $c^2>0$. Since $r=r(t)$ is positive throughout, (8) is non-oscillatory and has therefore both principal and non-principal solutions. To which of these two types will the particular solution $r=(x^2+y^2)^{\frac{1}{2}}$ belong? The answer turns out to be as follows: $r=(x^2+y^2)^{\frac{1}{2}}$ *is a non-principal or principal solution of* (8) *according as* (1) *is not or is non-oscillatory.*

In order to see this, consider the complex-valued solution $z=x+iy$ of (1). Then $z\neq 0$ and $z'\neq 0$, by (3), where $c\neq 0$, and it is also seen that the imaginary part of z'/z is c/r^2, where $r=|z|$. Hence $(\arg z)'=c/r^2$. This implies that $\arg(x+iy)$ is a strictly monotone function of t and that $1/r(t)$ is of class L^2 if and only if $\arg(x+iy)$ is bounded as $t\rightarrow\infty$. Since the latter condition is satisfied if and only if x (and/or y) fails to acquire an infinity of zeros as $t\rightarrow\infty$, it follows that $1/r(t)$ is of class L^2 if and only if (1) is non-oscillatory. This proves the last italicized statement.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] P. Hartman, "Differential equations with non-oscillatory eigenfunctions," *Duke Mathematical Journal*, vol. 15 (1948), pp. 697-709.

[2] ——— and A. Wintner, "Oscillatory and non-oscillatory linear differential equations," *American Journal of Mathematics*, vol. 71 (1949), pp. 627-649.

[3] ——— and A. Wintner, "On non-oscillatory linear differential equations," *ibid.*, vol. 75 (1953), pp. 717-730.

[4] A. Wintner, "Asymptotic integrations of the adiabatic oscillator in its hyperbolic range," *Duke Mathematical Journal*, vol. 15 (1948), pp. 55-67.

[5] ———, "On linear repulsive forces," *American Journal of Mathematics*, vol. 71 (1949), pp. 362-366.

[6] ———, "On almost free linear motions," *ibid.*, vol. 71 (1949), pp. 595-602.

MAJORANTS IN SPACES OF INTEGRABLE FUNCTIONS.*[1]

By G. G. Lorentz.

---

1. Introduction. Let $f(x)$ denote an integrable function on the interval $(0,1)$. The Hardy-Littlewood majorant of $f(x)$ is the function

$$\theta(x) = \theta(x;f) = \sup_{y} \int_{x}^{y} f(t)\,dt/(y-x), \tag{1}$$

provided it exists almost everywhere (a.e.). Hardy and Littlewood have shown (see for example [5], p. 244) that if $f \varepsilon L^p$, then $\theta(f)$ is defined and also belongs to L^p. More generally, if X is a certain space of integrable functions over $(0,1)$, we shall say that X has the *Hardy-Littlewood property* (or shortly that $X \varepsilon \mathrm{HLP}$) if $f \varepsilon X$ always implies $\theta(f) \varepsilon X$. The most general spaces X which we shall consider are the Köthe-Toeplitz spaces $X(C)$ defined as follows (compare [3]). Let C be some class of integrable non-negative functions $c(x)$ on $(0,1)$, which contains a function greater than a positive constant. A function $f(x)$ belongs to $X(C)$ and has norm $\| f \|$ if

$$\| f \| = \sup_{c \varepsilon C} \int_0^1 | f(x) | \, c(x)\,dx < +\infty. \tag{2}$$

$X(C)$ is a Banach space [3]; its dual X' is the space of all integrable functions g such that $\int_0^1 fg\,dx$ exists for all $f \varepsilon X$. In this case the suprema of $\int_0^1 fg\,dx$ for all f with $\| f \| \leqq 1$ and of $\int_0^1 f \, | g | \, dx$ for all $f \geqq 0$ with $\| f \| \leqq 1$ are equal and finite, and by definition equal to the norm of g in X'. Thus our "dual space of X" is different from "conjugate space of X" in Banach's terminology (i. e. the space of all bounded linear functionals on X). The use of dual spaces rather than conjugate spaces is justified by the fact that the former are simpler and that the second dual of X may be X even if X is not reflexive. Actually, the relation $X'' = X$ holds for any Köthe-Toeplitz space.

The importance of Köthe-Toeplitz spaces is underlined by the fact that they are identical (as has been shown by the author in a paper under prepara-

---

* Received December 2, 1954.

[1] This work was supported by the research grant NSF G 1014 at Wayne University, sponsored by the National Science Foundation.

tion for publication) to an apparently more general class of spaces ([2], p. 415) described in terms of lattice operations in X and the norm $\| f \|$ only, without reference to the class $C = \{c\}$. Essentially identical with these are spaces discussed by Ellis and Halperin [6], p. 576.

Our problem will be to give necessary and sufficient conditions for C in order that a Köthe-Toeplitz space $X(C)$ should have the Hardy-Littlewood property.

We shall assume that C is *rearrangement-invariant,* that is, with each $c(x)$ any of its rearrangements [3] $c_r(x)$ also belong to C. (Then $X(C)$ will also be rearrangement-invariant.) In this case the answer is easily given (see Section 2): for each $c \varepsilon C$, the function

$$\bar{c}(x) = \int_x^1 v(t)/t\,dt \tag{3}$$

should belong to X'.

Special cases of spaces $X(C)$ with rearrangement-invariant C are spaces $\Lambda(\phi)$, $\mathrm{M}(\phi)$ [3] and Orlicz spaces L_Φ (for an exposition of the theory of Orlicz spaces see [4]). Here the condition just given may be simplified so as to bear directly on ϕ or Φ. This is carried out in Sections 3-5. Spaces $\Lambda(\phi, p)$ and $\mathrm{M}(\phi, p)$ (see [3]), with $p > 1$, are not very interesting in our problem, since the original argument of Hardy and Littlewood is sufficient to show ([1], [2]) that they have the Hardy-Littlewood property.

The importance of the Hardy-Littlewood property lies in the fact that it is equivalent to the *dominated convergence* a. e. of the quotients $f_h(x) = h^{-1}\{F(x+h) - F(x)\}$, where $F(x)$ is the indefinite integral of $f(x)$, to the latter function, whenever $f(x)$ lies in X, the common majorant for the f_h being $\theta(x;f)$. In lattice-theoretic terms, this dominated convergence may be shown to be equivalent to the *order convergence* of f_h to f in the Banach lattice X endowed with its natural partial order. The dominated convergence of f_h to f implies dominated convergence to f of many types of singular integrals; for example, the Fejér and Poisson integrals (compare [5], p. 247). These facts justify our interest in the Hardy-Littlewood property.

2. Hardy-Littlewood property of the spaces $X(C)$. For a positive and decreasing function $f(x)$, the majorant $\theta(x;f)$ is obviously equal to the function

$$\bar{f}(x) = \int_0^x f(t)\,dt/x. \tag{4}$$

It is therefore natural to consider the properties of a space $X(C)$, which we call HLP′ and HLP″, defined by the requirements that $\bar{f} \varepsilon X(C)$ for all $f \varepsilon X(C)$ or that $\bar{f} \varepsilon X(C)$ for all positive decreasing f in $X(C)$.

Since $\bar{f}(x)$ is contained between $\theta(x;f)$ and

$$-\theta(x;-f) = \inf_{y} \int_{y}^{x} f(t)\,dt/(y-x),$$

it is clear that HLP implies HLP′, and this in turn implies HLP″.

Lemma. *If C is rearrangement-invariant, then all properties* HLP, HLP′, HLP″ *are equivalent, and each implies*

(5) $\quad \| \theta(f) \| \leqq A \| f \|$ *and* (6) $\quad \| \bar{f} \| \leqq A \| f \|$

with some constant A.

Proof. If HLP″ holds and $f \varepsilon X(C)$, then f^*, the decreasing rearrangement of $|f|$, also belongs to $X(C)$, and hence $\bar{f}^* = \theta(f^*) \varepsilon X(C)$. By [2], pages 420-421, $\theta^*(x;f) \prec 2\theta(x;f^*)$. (This means that

$$\int_0^x \theta^*(t;f)\,dt \leqq 2 \int_0^x \theta(t;f^*)\,dt$$

for all $x \geqq 0$.) Therefore

$$\int_0^1 c\theta\,dx \leqq \int_0^1 c^*\theta^* dx \leqq 2 \int_0^1 c^*(x)\theta(x;f^*)\,dx \leqq 2 \; \| \theta(f^*) \|,$$

so that $\theta(f) \varepsilon X(C)$, and we have established HLP. Also, if (6) is satisfied for all positive decreasing f, then (5) must hold. It will now be sufficient to show that HLP implies (6). Let $T_n f = \bar{f}_n$ be the sequence of operators from $X(C)$ into itself, defined by placing $\bar{f}_n(x) = \bar{f}(x)$ or $\bar{f}_n(x) = 0$ according as $1/n \leqq x \leqq 1$ or $0 \leqq x < 1/n$. By [3], Theorem 3.8.3, T_n are bounded linear operators. For $x > 1/n$ we have $-\theta(x;-f) \leqq \bar{f}_n(x) \leqq \theta(x;f)$, hence there is a function F in $X(C)$, e.g., $F = |\theta(f)| + |\theta(-f)|$, so that $|\bar{f}_n(x)| \leqq F(x)$ for all x and n. Also, $\bar{f}_n(x) \rightarrow \bar{f}(x)$ as $n \rightarrow \infty$. This implies that $\bar{f}_n \rightarrow \bar{f}$ in the norm of the space $X(C)$. Since the limit of a convergent sequence of linear bounded operators is bounded, we see that the operator $Tf = \bar{f}$ has a bound. This establishes (6) and completes the proof.

For $f \varepsilon X(C)$, $f \geqq 0$ we have, with the notation (3),

$$\begin{aligned} \| f \|_X &= \sup_{c \varepsilon C} \int_0^1 c(x)\,dx/x \int_0^x f(t)\,dt \\ &= \sup_{c \varepsilon C} \int_0^1 f(x)\,dx \int_x^1 c(t)/t\,dt = \sup_{c \varepsilon C} \int_0^1 f(x)\bar{c}(x)\,dx. \end{aligned} \tag{7}$$

The necessary and sufficient condition for $X(C) \varepsilon \mathrm{HLP}$ is, by the Lemma, that $\| \bar{f} \| \leqq A \| f \|$, $f \varepsilon X(C)$. Remembering the definition of the dual space, we see that $X(C) \varepsilon \mathrm{HLP}$ is equivalent to $\| \bar{c} \|_{X'} \leqq A$. In this argument we could have restricted $f \varepsilon X(C)$ to be positive and decreasing. Let C^* denote the subclass of C consisting of all positive decreasing $c^* \varepsilon C$. Then instead of (7) we get, since $\bar{f}(x)$ is positive and decreasing,

$$\| \bar{f} \|_X = \sup_{c^* \varepsilon C^*} \int_0^1 c^*(x) dx/x \int_0^x f(t) dt = \sup_{c \varepsilon C^*} \int_0^1 f(x) \bar{c}(x) dx.$$

Hence we obtain:

THEOREM 1. *If C is rearrangement-invariant, then $X(C) \varepsilon \mathrm{HLP}$ if and only if for some constant A,*

$$\| \bar{c} \|_{X'} \leqq A \tag{8}$$

for all $c \varepsilon C$, or, equivalently, for all $c \varepsilon C^*$.

We note also another condition for $X(C) \varepsilon \mathrm{HLP}$. For $0 \leqq a \leqq 1$, let f_a denote the function $f_a(x) = f(ax)$. Since

$$\| \bar{f} \| = \sup_{c \varepsilon C} \int_0^1 c(x) dx/x \left| \int_0^x f(t) dt \right| \leqq \sup_c \int_0^1 c(x) dx \int_0^1 | f(ax) | \, da$$
$$\leqq \int_0^1 \sup_c \int_0^1 | f(ax) | \, c(x) dx \, da = \int_0^1 \| f_a \|_X \, da,$$

we obtain by Lemma, that for $X(C) \varepsilon \mathrm{HLP}$ it is *sufficient* that for some constant A and all $f \varepsilon X(C)$,

$$\int_0^1 \| f_a \|_X \, da \leqq A \| f \|_X. \tag{9}$$

3. Spaces $\Lambda(\varphi)$ and $M(\varphi)$. Let $\phi(x)$ be a decreasing positive integrable function on $(0,1)$, which we shall assume zero for $x > 1$, and let $\Phi(x) = \int_0^x \phi(t) dt$. The space $\Lambda(\phi)$ consists of all functions $f(x)$ such that

$$\| f \|_\Lambda = \int_0^1 \phi(x) f^*(x) dx = \sup_r \int_0^1 \phi_r(x) | f(x) | \, dx < +\infty; \tag{10}$$

the supremum is taken over all rearrangements ϕ_r of ϕ. Dual with $\Lambda(\phi)$ is the space $M(\phi)$ of functions f with

$$\| f \|_M = \sup_a (1/\Phi(a)) \int_0^a f^*(t) dt < +\infty. \tag{11}$$

Obviously $\Lambda(\phi)$ and $M(\phi)$ are spaces $X(C)$ with rearrangement-invariant C.

For $\Lambda(\phi)$, C^* consists of the single function ϕ. Applying Theorem 1 and formula (11) to this C^*, we see that $\Lambda(\phi) \varepsilon \mathrm{HLP}$ exactly if

$$\int_0^a dx \int_x^1 \phi(t)/t\,dt \leqq A \int_0^a \phi(t)\,dt, \qquad 0 \leqq a \leqq 1.$$

Changing the order of integrations we see that this is equivalent to the existence of a constant M such that

$$\int_a^1 \phi(x)/x\,dx \leqq Ma^{-1} \int_0^a \phi(x)\,dx, \qquad 0 \leqq a < 1. \tag{12}$$

This condition will be simplified.

THEOREM 2. *A space $\Lambda(\phi)$ has the Hardy-Littlewood property if and only if*

$$\limsup_{a \to 0} \Phi(2a)/\Phi(a) < 2. \tag{13}$$

Proof. If (13) holds, there is an a_0 and an $\epsilon > 0$ such that $\Phi(2a)/\Phi(a) < 2 - \epsilon$ for $0 < a \leqq a_0$. This quotient is a continuous function of a for $a \geqq a_0$, which can assume the value 2 only if $\phi(x)$ is constant, which is ruled out by (13). Hence we can assume that the inequality holds for all $0 < a \leqq 1$. Then we shall have

$$\int_a^{2a} \phi(x)\,dx \leqq (1-\epsilon) \int_0^a \phi(x)\,dx, \qquad 0 < a \leqq 1, \tag{14}$$

$$\int_0^{2^k a} \phi(x)\,dx \leqq 2^k(1-\epsilon/2)^k \int_0^a \phi(x)\,dx. \tag{15}$$

These two inequalities give

$$\begin{aligned}
\int_a^1 \phi(x)/x\,dx &\leqq \int_0^1 \phi(x)dx/(x+a) = \int_0^a + \int_a^{2a} + \cdots \\
&\qquad + \int_{2^{k-1}a}^{2^k a} \phi dx/(x+a) + \cdots \\
&\leqq 1/a \int_0^a \phi dx + (1-\epsilon)/a \int_0^a \phi dx + \cdots + (1-\epsilon)/(2^{k-1}a) \int_0^{2^{k-1}a} \phi dx + \cdots \\
&\leqq \{1 + (1-\epsilon) + (1-\epsilon)(1-\epsilon/2) + \cdots \\
&\qquad + (1-\epsilon)(1-\epsilon/2)^{k-1} + \cdots\}(1/a) \int_0^a \phi dx \\
&= (M/a) \int_0^a \phi dx,
\end{aligned}$$

with $M = (2-\epsilon)/\epsilon$. This proves that (13) is sufficient.

Now assume that (13) does not hold. We shall show that (12) is false for an arbitrarily large M.

For M chosen, take a natural number $n > 2M$ and an $\epsilon > 0$ so small that

$$(16) \qquad \tfrac{1}{2}(1-2^n\epsilon)n > M.$$

By hypothesis, there is a b with $0 < b < 1$ such that

$$(17) \qquad \int_0^{2b}\phi dx \geqq (2-\epsilon)\int_0^{b}\phi dx, \qquad \int_b^{2b}\phi dx \geqq (1-\epsilon)\int_0^{b}\phi dx.$$

Then

$$(18) \qquad \int_{b/2}^{b}\phi dx \geqq (1-2\epsilon)\int_0^{b/2}\phi dx,$$

for otherwise we would have

$$\int_b^{2b}\phi dx \leqq 2\int_{b/2}^{b} < \int_{b/2}^{b} + (1-2\epsilon)\int_0^{b/2} \leqq (1-\epsilon)\int_0^{b}\phi dx,$$

which contradicts (17). Repeating this argument several times, we obtain for $k = 1, 2, \cdots, n$,

$$(19) \qquad \int_{b/2^k}^{b/2^{k-1}}\phi dx \geqq (1-2^k\epsilon)\int_0^{b/2^k}\phi dx \geqq (1-\delta)\int_0^{b/2^k}\phi dx,$$

where $\delta = 2^n\epsilon$. Furthermore,

$$\int_0^{b/2^k}\phi(x)dx \geqq (2-\delta)\int_0^{b/2^{k+1}}\phi dx \geqq \cdots \geqq (2-\delta)^{n-k}\int_0^{b/2^n}\phi dx.$$

After these preparations we obtain for our special choice of b, by (18) and (19),

$$\begin{aligned}\int_{b/2^n}^{1}\phi(x)/x\,dx &\geqq 2^{n-1}/b\int_{b/2^n}^{b/2^{n-1}}\phi dx + 2^{n-2}/b\int_{b/2^{n-1}}^{b/2^{n-2}} + \cdots + 1/b\int_{b/2}^{b} \\ &\geqq (1-\delta)/b\{2^{n-1}\int_0^{b/2^n} + 2^{n-2}\int_0^{b/2^{n-1}} + \cdots + \int_0^{b/2}\} \\ &\geqq 2^{n-1}(1-\delta)/b\{\int_0^{b/2^n} + (1-\delta/2)\int_0^{b/2^n} + \cdots + (1-\delta/2)^{n-1}\int_0^{b/2^n}\} \\ &\geqq \tfrac{1}{2}(1-\delta)^n n\, 2^n/b\int_0^{b/2^n}\phi dx.\end{aligned}$$

Because of (16), we see that for the chosen M, (12) is violated for $a = b/2^n < 1$. Thus (13) is necessary.

THEOREM 3. *A space* M(ϕ) *has the Hardy-Littlewood property if and only if, for some constant A,*

$$(20) \qquad \int_0^{a}\Phi(x)/x\,dx \leqq A\Phi(a), \text{ for } 0 \leqq a \leqq 1,$$

Proof. The functions $c(x)$ of the class C^* are here $c(x)=\Phi(a)^{-1}$ or $c(x)=0$ according as $0\leqq x\leqq a$ or $a<x\leqq 1$. Since the dual to $\mathrm{M}(\phi)$ is $\Lambda(\phi)$, the condition (8) becomes

$$\int_0^1 \phi(x)\int_x^1 c(t)/t\,dt\,dx \leqq A,$$

and this is equivalent to (20).

4. Orlicz spaces. Let $0\leqq\phi(u)\leqq+\infty$ be a function defined for $u\geqq 0$, increasing to infinity with u and vanishing for $u=0$, but not identically 0 or $+\infty$ for all $u>0$. Let $\psi(u)$ be its inverse, defined in a familiar way everywhere, except perhaps on a countable set of points corresponding to the discontinuities of ϕ. Let $\Phi(x)$ and $\Psi(x)$ be integrals of ϕ and ψ over $(0,x)$; Φ and Ψ are called conjugate functions (in sense of W. H. Young). The Orlicz space L_Φ consists of all functions $f(x)$ on $(0,1)$ such that $\int_0^1 fg\,dx$ exists for all g with

$$I_\Psi(g)=\int_0^1 \Psi(|g|)\,dx<+\infty,$$

and we define

$$\|f\|_\Phi=\sup_{I_\Psi(g)\leqq 1}\int_0^1 fg\,dx.$$

This space L_Φ is a Banach space; it contains all functions f with $I_\Phi(f)<+\infty$, but in general also other functions. For example, if $\phi(u)=0$ for $0\leqq u\leqq 1$ and $\phi(u)=+\infty$ for $u>1$, L_Φ is the space M of bounded functions, but $I_\Phi(f)<+\infty$ means that $|f(x)|\leqq 1$ a. e. However, L_Φ may be described as the set of all functions f such that for some $\lambda>0$, $I_\Phi(\lambda f)<+\infty$; indeed, it may be shown [4] that $I_\Phi(f/\|f\|_\Phi)\leqq 1$. It follows that the spaces L_Φ, L_Ψ are dual to each other. Under the additional hypothesis that $\Phi(2u)/\Phi(u)$ is bounded, a stronger result is true: L_Ψ is the conjugate space to L_Φ. For the information of the reader we note that the boundedness of $\Phi(2u)/\Phi(u)$ is also necessary and sufficient for the space L_Φ to be separable. While the sufficiency of the condition is well known, its necessity seems to have escaped the attention of the writers on this subject until recently, when it was discovered independently by Mr. W. Luxemburg (Delft), Sobolev and Krasnosel'skiĭ, and the author.

THEOREM 4. *A space L_Φ has the Hardy-Littlewood property if and only if the conjugate function Ψ satisfies*

$$\Psi(2b)\leqq M\Psi(b),\qquad b\geqq b_0, \tag{21}$$

for some constants M and $b_0>0$.

Proof. It is easy to see that (21) implies the existence of two positive constants α, C such that

$$\Psi(ab) \leqq Ca^{\alpha}\Psi(b), \qquad b \geqq b_0, a > 1. \tag{22}$$

To prove that the condition is sufficient we shall show that it implies (9). For $f_a(x) = f(ax)$, $f \varepsilon L_\Phi$ we have with $c(x) \geqq 0$,

$$\| f_a \|_\Phi = \sup_{I_\Psi(c) \leqq 1} \int_0^1 | f(ax) | \, c(x) \, dx.$$

We shall begin by replacing arbitrary positive functions $c(x)$ by others such that, for each x, either $c(x) \geqq b_0$ of $c(x) = 0$. To achieve this, assume that $c(x) \geqq 0$ and $I_\Psi(c) \leqq 1$, and decompose the interval $(0, 1)$ into two sets E_1, E_2 defined by the inequalities $c(x) < b_0$, and $c(x) \geqq b_0$. Since

$$\int_{E_1} | f(ax) | \, dx \leqq \int_0^1 | f(x) | \, dx \leqq A \, \| f \|_\Phi$$

(the last inequality follows from the fact that a small constant $c(x) = \lambda > 0$ satisfies $I_\Psi(c) \leqq 1$), we have

$$\| f_a \|_\Phi \leqq A \, \| f \|_\Phi + \sup_{I_\Psi(c') \leqq 1} \int_0^1 | f(ax) | \, c'(x) \, dx,$$

where the $c'(x)$ are such that either $c'(x) \geqq b_0$, or $c'(x) = 0$. On substitution $ax = t$ we obtain that the last supremum is

$$\begin{aligned} &\sup_{\int_0^a \Psi(c')dt \leqq a} \cdot 1/a \int_0^a | f(t) | \, c'(t) \, dt \\ &\qquad \leqq \sup_{I_\Psi(c') \leqq a} 1/a \int_0^a | f(t) | \, c'(t) \, dt. \end{aligned} \tag{23}$$

Assume first that $Ca \leqq 1$. Because of (22), $I_\Psi(c') \leqq a$ implies $I_\Psi((Ca)^{-1/\alpha}c') \leqq 1$. Applying this to (23) and replacing $(Ca)^{-1/\alpha}c'(t)$ by $c(t)$, we obtain in this case

$$\begin{aligned} \| f_a \|_\Phi &\leqq A \, \| f \|_\Phi + \sup_{I_\Psi(c) \leqq 1} 1/a(Ca)^{1/\alpha} \int_0^1 | f | c \, dt \\ &= (C_1 a^{(1-\alpha)/\alpha} + A) \, \| f \|_\Phi. \end{aligned}$$

For $Ca \geqq 1$, (23) readily gives $\| f_a \| \leqq C_2 \, \| f \|_\Phi$, and we obtain the condition (9).

We shall prove the necessity of (21) by means of Theorem 1. Assume that (21) does not hold. Taking into account the properties of Orlicz spaces, it will be sufficient to show that there is a function $c(x) \geqq 0$ which belongs to L_Ψ and such that $I_\Psi(\lambda\bar{c}) = +\infty$ for each $\lambda > 0$; this will imply that $\bar{c} \notin L_\Psi$. If (21) does not hold and $\Psi(u)$ is finite for all $0 \leqq u < +\infty$, there must exist a sequence $b_k \to \infty$ such that $\Psi(2b_k)/\Psi(b_k) \to \infty$. Let the integers n_k be defined by

$$2^{-k-1} < 2^{-n_k}\Psi(b_k) \leqq 2^{-k}, \qquad k = 1, 2, \cdots. \tag{24}$$

We may assume (passing, if necessary, to a subsequence of the b_k) that $n_{k+1} > n_k + k$ and that $\Psi(2b_k)/\Psi(b_k) \geqq 2^{2k+1}$, so that

$$2^{-n_k}\Psi(2b_k) \geqq 2^k. \tag{25}$$

Now we define $c(x) = b_k$ for $2^{-n_k-k} \leqq x < 2^{-n_k}$, $k = 1, 2, \cdots$, and $c(x) = 0$ elsewhere. Then, by (24),

$$\int_0^1 \Psi(c)\,dx \leqq \sum 2^{-n_k}\Psi(b_k) \leqq \sum 2^{-k} < +\infty,$$

and therefore $c \in L_\Psi$. On the other hand, for each $\lambda > 0$ we have by definition of $\bar{c}$,

$$\int_0^1 \Psi(\lambda\bar{c})\,dx \geqq \sum_{k=1}^{\infty} \int_{2^{-n_k-k}}^{2^{-n_k-k+1}} \Psi\Big(\lambda \int_{2^{-n_k-k+1}}^{2^{-n_k}} b_k/t\,dt\Big)\,dx$$

$$= \sum_{k=1}^{\infty} 2^{-n_k-k}\Psi(\lambda b_k(k-1)\log 2).$$

For all sufficiently large values of k, say for $k \geqq k_0 = k_0(\lambda)$, the argument of Ψ in the last sum is greater than $2b_k$, and by (25), the series diverges. Hence $\int_0^1 \Psi(\lambda\bar{c})\,dx = +\infty$.

In the case when (21) is violated in such a way that $\Psi(u)$ becomes infinite for finite u, the proof is similar, but somewhat simpler.

WAYNE UNIVERSITY.

REFERENCES.

[1] P. L. Butzer, "On Bernstein Polynomials," Thesis, University of Toronto (1951).

[2] G. G. Lorentz, "On the theory of spaces Λ," *Pacific Journal of Mathematics*, vol. 1 (1951), pp. 411-429.

[3] ———, *Bernstein Polynomials*, Toronto, 1953.

[4] A. C. Zaanen, *Linear Analysis*, New York and Amsterdam, 1953.

[5] A. Zygmund, *Trigonometric Series*, Warszawa-Lwów, 1935.

[6] H. W. Ellis and I. Halperin, "Function spaces," *Canadian Journal of Mathematics*, vol. 5 (1953), pp. 576-592.

ON ALGEBRAIC GROUPS AND HOMOGENEOUS SPACES *

By André Weil.

In a recent paper in the same Journal ([4] of the bibliography; quoted hereafter as AG), I gave some results on algebraic groups and transformation-spaces, which supplement those in my *Variétés abéliennes* ([3]; quoted as VA). Applications will now be made of that theory to somewhat more specific questions. In no. 1, a rather general procedure is described for obtaining, from a given transformation-space S with respect to a group G and from a suitable cycle on S, another transformation-space with respect to the same group. As shown in no. 2, this includes as a special case the construction of coset-spaces and of factor-groups; thanks to the main theorem in AG, these can now be defined without enlarging the groundfield, whereas such an enlargement was required in their construction as previously given by S. Nakano ([2]); except for this, we have substantially followed his method.

The rest of this paper is chiefly devoted to "principal homogeneous spaces," i. e. to those homogeneous spaces on which the group operates in a simply transitive manner. The pair consisting of such a space and of one point on it does not differ materially from a group; thus there is little incentive for studying those spaces as long as one is not paying any attention to the groundfield or if the groundfield is algebraically closed. But it can happen that a principal homogeneous space contains no rational point over the groundfield over which it is defined; an example of this is given by the plane curve $X^3+pY^3=p^2$ over the rational number-field, where p is a rational prime; this may be considered as a principal homogeneous space with respect to its jacobian variety, which is the plane curve $Y^2=4X^3-27$. More generally, Chow's work (cf. [1]) has shown that it is not always possible to map a curve "canonically" into its jacobian variety by a mapping defined over the groundfield, but that the curve can always be so mapped into a suitable principal homogeneous space with respect to its jacobian variety. This, among other results, will be proved again here by a different method, which can be extended at once to a variety V of higher dimension and to its Albanese variety, provided the groundfield is one over which the latter is defined. It will also be shown that the classes of principal homogeneous

---

* Received December 27, 1954.

spaces with respect to a commutative group can be arranged into a torsion-group, i. e. a group whose elements are all of finite order; and it follows at once from the results of no. 5 that this group must be countable if the groundfield is finitely generated over the prime field. There seems to be no reason why it should be finite, even if the groundfield is the field of rational numbers; a more detailed investigation of its structure, e. g. for the case of an elliptic curve over the field of rational numbers, would be of considerable interest from the point of view of the theory of diophantine equations.

1. Let G be a group and S a transformation-space with respect to G, both defined over a field k. Let Z be either a divisor on S or a cycle on S whose components have coefficients which are prime to the characteristic of the universal domain. We denote by sZ, for any $s \epsilon G$, the transform of Z by the mapping $u \rightarrow su$ of S onto itself. Let K be an overfield of k over which Z is rational. Let x be generic over K on G; by Prop. 6 of the Appendix of AG, there is a finitely generated extension $k(t)$ of k which is the smallest overfield of k over which xZ is rational; as xZ is rational over $K(x)$, we have $k(t) \subset K(x)$. If x' is also a generic point of G over K, and σ is the isomorphism of $K(x)$ onto $K(x')$ over K which maps x onto x', σ will transform xZ into $x'Z$; if t' is the image of t under σ, $k(t')$ will then be the smallest overfield of k over which $x'Z$ is rational, and, by Prop. 6 of the Appendix of AG, the cycle $x'Z$ depends only upon t'; in other words, if σ_1 is an isomorphism of $K(x)$ onto a field $K(x_1')$, mapping x onto x_1' and t onto t_1', we have $x'Z = x_1'Z$ if and only if $t' = t_1'$. In particular, take $x' = yx$, with y generic over $K(x)$ on G; then $k(t')$ is the smallest overfield of k over which yxZ is rational; as yxZ can be written as $y(xZ)$, it is rational over $k(y, t)$, so that $k(t') \subset k(y, t)$; similarly, xZ can be written as $y^{-1}(yxZ)$ and is therefore rational over $k(y, t')$, so that $k(t) \subset k(y, t')$. This shows that $k(y, t) = k(y, t')$.

Put now $z = yx$, so that we have $k(t') \subset K(z)$; take y' generic over $K(x, y)$ on G, and call τ the isomorphism of $K(z)$ onto $K(y'z)$ over K which maps z onto $y'z = y'yx$; let t'' be the image of t' under τ. Then t'' is the image of t under the isomorphism $\tau \circ \sigma$ of $K(x)$ onto $K(y'yx)$ over K which maps x onto $y'yx$.

If Z is such that $k(t)$ is a regular extension of k, then we may call T the locus of t over k and, with the above notations, we may write $t' = g(y, t)$, where g is a mapping of $G \times T$ into T, defined over k. Then the results we have just proved mean that g satisfies (TG 1, 2) of AG, no. 2, i. e. that it is a normal law between G and T. Applying now the main theorem of AG, we get the following result:

PROPOSITION 1. *Let G be a group and S a transformation-space with respect to G, both defined over a field k. Let Z be either a divisor on S or a cycle on S whose components have coefficients which are prime to the characteristic. Let K be an overfield of k over which Z is rational; and assume that, if x is generic over K on G, the smallest extension k' of k over which xZ is rational is a regular extension of k. Then there is a transformation-space T with respect to G, defined over k, and an everywhere defined mapping F of G into T, defined over K, such that the point $t = F(x)$ is generic over k on T, that $k' = k(t)$, and that $F(ss') = sF(s')$ for all s, s' on G. For any s, s' in G, we have $F(s) = F(s')$ if and only if $sZ = s'Z$. If Z is algebraic over k, one can take for T a homogeneous space with respect to G.*

The existence of a transformation-space T with a generic point t over k such that $k' = k(t)$ has been proved above; moreover, with the same notations as above, we have $k(t) \subset K(x)$, $k(t') \subset K(yx)$, $t' = yt$; as $K(x)$, $K(yx)$ are independent extensions of K, this shows, if K is algebraic over k, that $k(t)$, $k(t')$ are then independent extensions of k, i.e. that T is pre-homogeneous, so that, by the main theorem of AG, we may replace it by a birationally equivalent homogeneous space. As $k(t) \subset K(x)$, we may write $t = F(x)$, with F defined over K; then we have $yt = F(yx)$, i.e. $F(yx) = yF(x)$, for y generic over $K(x)$ on G. This may be written as $F(x) = y^{-1}F(yx)$, which shows, if s is any point of G and y is taken generic over $K(s)$ on G, that F is defined at s. As F is everywhere defined, the relation $F(yx) = yF(x)$ implies $F(ss') = sF(s')$ for all s, s' on G. Now, for any s, s' on G, take x generic over $K(s, s')$ on G; then xs, xs' are both generic over K on G, and therefore, if σ is the isomorphism of $K(xs)$ onto $K(xs')$ over K which maps xs onto xs', σ will map $F(xs)$ onto $F(xs')$ and the cycle xsZ onto $xs'Z$; then, as we have seen above, we have $xsZ = xs'Z$ if and only if $F(xs) = F(xs')$; as the latter relation can be written $xF(s) = xF(s')$, so that these two relations are respectively equivalent to $sZ = s'Z$ and to $F(s) = F(s')$, this completes our proof.

2. We first apply Prop. 1 to the construction of the homogeneous space defined by a group G and a subgroup of G.

PROPOSITION 2. *Let G be a group, defined over a field k; let Z be a rational cycle over k on G, consisting of components with coefficient* 1, *and such that its support $|Z|$ is a subgroup of G. Then there is a homogeneous space H with respect to G, defined over k, and a rational point a over k on H, with the following properties:* (i) *if we put, for a generic x over k on G,*

$F(x) = xa$, the mapping F of G onto H determines a one-to-one mapping of the cosets of $|Z|$ in G onto the points of H; (ii) *$k(x)$ is separable over $k(F(x))$;* (iii) *if ϕ is a mapping of G into a variety V, defined over an overfield K of k, and such that $\phi(xs) = \phi(x)$ whenever $s \in |Z|$ and x is generic over $K(s)$ on G, there is a mapping ψ of H into V, defined over K, such that $\phi = \psi \circ F$. If $|Z|$ is a normal subgroup of G, then one can define on H a group-law, defined over k, such that F is a homomorphism of G onto H.*

The "support" of a cycle was defined at the beginning of the Appendix of AG. The assumption on $|Z|$ implies that Z has one and only one component Z_0 containing e, that this is a subgroup of G, and that all the components of Z are cosets of Z_0 in G. We apply Prop. 1 to the cycle Z on $S = G$, G acting on itself by the left-translations, and to the field k; if x is generic over k on G, the smallest extension of k over which xZ is rational is contained in $k(x)$, and hence, by AG-App., Prop. 3, Coroll., it is regular over k. Therefore, by Prop. 1, there is a homogeneous space with respect to G, which we now call H, defined over k, and a mapping F of G into H, defined over k, with the properties stated in Prop. 1; in particular, F is everywhere defined, $t = F(x)$ is generic over k on H, and $k(t)$ is the smallest extension of k over which xZ is rational. If we put $a = F(e)$, a is rational over k, and we have, for all $s \in G$, $F(s) = sa$. If s, s' are any two points on G, we have $sa = s'a$ if and only if $sZ = s'Z$, i. e. if and only if $s^{-1}s'Z = Z$; by the assumption on Z, the latter relation is equivalent to $s^{-1}s' \in |Z|$. Thus the points of H are in a one-to-one correspondence with the cosets of $|Z|$ in G.

Call Γ the graph of F on $G \times H$. For any $b \in H$, $\Gamma \cap (G \times b)$ is the set of those points (s, b) which are such that $sa = b$; in particular, if x is generic over k on G and if we put $t = F(x) = xa$, $\Gamma \cap (G \times t)$ is the set of the points (s, t) such that $sa = xa$, i. e. $x^{-1}s \in |Z|$; this set can be written as $|xZ| \times t$. As $\Gamma \cdot (G \times t)$ is the prime rational cycle over $k(t)$ on $G \times H$ with the generic point (x, t) over $k(t)$, this shows that the prime rational cycle Z' over $k(t)$ on G with the generic point x has the same components as the cycle xZ; as the latter is rational over $k(t)$ and its components have the coefficient 1, this implies that $Z' = xZ$. As the components of the prime rational cycle with the generic point x over $k(t)$ have the coefficient 1, $k(x)$ must be separable (i. e. "separably generated") over $k(t)$.

As to (iii), let x be generic over K on G; put $t = F(x)$ and $w = \phi(x)$; let w' be any generic specialization of w over $K(t)$; this can be extended to a generic specialization x' of x over $K(t)$; we have then $w' = \phi(x')$

and $t = F(x')$, and the latter relation implies that x' is on $|xZ|$, i.e. that it is of the form xs with $s \in |Z|$. Let $\bar{x}$ be generic on G over $K(s)$; we have $\phi(\bar{x}s) = \phi(\bar{x})$; specializing $\bar{x}$ to x over $K(s)$, we get $\phi(xs) = \phi(x)$, since both sides are defined, i.e. $w' = w$. This shows that w is purely inseparable over $K(t)$; as it is at the same time rational over $K(x)$ which is separable over $K(t)$, it is therefore rational over $K(t)$, and we may write $w = \psi(t)$, where ψ is a mapping of H into V, defined over K. This proves (iii).

Finally, assume that $|Z|$ is a normal subgroup of G; let x, y be independent generic points of G over k; put $t = F(x)$, $u = F(y)$. We have $F(xy) = xF(y) = xu$; this is a function of x, defined over $k(u)$. If $s \in |Z|$, we have $xsy = xys'$ with $s' = y^{-1}sy \in |Z|$, and therefore $F(xsy) = F(xy)$; by (iii) applied to the mapping $x \rightarrow xu$ of G into H and to $K = k(u)$, this implies that $F(xy)$ is rational over $k(u, t)$. Therefore the mapping $u \rightarrow xu$ of H into H is defined over $k(t)$; on the other hand, if k' is any field of definition for that mapping, containing k, the image $xa = t$ of a by it is rational over k', so that $k(t) \subset k'$; thus $k(t)$ is the smallest field of definition for the mapping $u \rightarrow xu$. This shows that G is not operating faithfully on H; applying Prop. 2 of AG, no. 3, to G and H, we see that we can define on H a normal law of composition f such that $F(xy) = f(F(x), F(y))$. By the main theorem of AG, we can then replace H by a birationally equivalent group H', defined over k, with a mapping F' of G into H', also defined over k, such that $F'(xy) = F'(x)F'(y)$, $F(x)$ and $F'(x)$ being corresponding generic points of H and H' over k when x is generic over k on G. As usual, from the relation $F'(x) = F'(y)^{-1}F'(yx)$ which holds for x, y generic and independent over k, we deduce that F' is everywhere defined[1]; therefore, if G is made to operate on H' by the law $(x, w) \rightarrow F'(x)w$ for x, w generic and independent over k on G and H', H' is a homogeneous space with respect to G. But then the unicity assertion in the main theorem of AG can be applied to H and H' and shows that they are biregularly equivalent; in other words, H itself, with the law f, is a group. This completes the proof.

It is easily seen that the pair (H, a) is uniquely determined, up to an isomorphism, by the conditions (i), (ii), (iii) in Prop. 2; in other words, if H' and a' have similar properties, there is an everywhere biregular birational correspondence between H and H' which maps a onto a' and transforms the law of composition between G and H into the law between G and H'. The space H may be called the *coset-space* determined by G and Z, and may be denoted by G/Z; if $|Z|$ is a normal subgroup of G, the space H, with the group-law determined by Prop. 2, is called the *quotient-group* (or factor-group) of G by Z, and is denoted by G/Z.

---

[1] This is Theorem 1 of Nakano ([2]).

3. Before making another application of Prop. 1, we will introduce a new condition which a law of composition may satisfy. Let V, W be two varieties, g a mapping of $V \times W$ into W, and k a field of definition for V, W and g; consider the following condition:

(TG 1′) *If x, u are independent generic points of V, W over k, and $v = g(x,u)$, then $k(x,u) = k(x,v) = k(u,v)$.*

The condition $k(x,u) = k(x,v)$ is equivalent to (TG 1) of AG, no. 2. The condition $k(x,u) = k(u,v)$ implies that the dimension of V, which is the dimension of x over $k(u)$, is the same as that of v over $k(u)$, and therefore at most that of W; if the dimensions of V and of W are the same, this implies that v is generic over $k(u)$ on W, which is condition (H) of AG, no. 2. Let k' be any field of definition containing k for the birational correspondence $u \rightarrow v = g(x,u)$ between W and itself; if u is taken generic over $k'(x)$ on W, we have $k'(u) = k'(v)$; since $k(x) \subset k'(u,v)$ by (TG 1′), we have $k(x) \subset k'(u)$. Taking u' generic on W over $k'(x,u)$, we get in the same manner $k(x) \subset k'(u')$. As $k'(u)$, $k'(u')$ are independent regular extensions of k', their intersection is k', so that $k(x) \subset k'$. This shows that (TG 1′) implies (TG 3). In view of the results of AG, end of no. 3, this shows that, if g satisfies (TG 1′) and (TG 2′), or if two mappings f, g of $V \times V$ into V and of $V \times W$ into W are given and satisfy (TG 1′. 2), then V is a pre-group and W a pre-transformation space, and V operates faithfully on W.

If a pre-group V and a pre-transformation space W satisfy (TG 1′), we say that W is a *pre-principal space* with respect to V; if at the same time V and W have the same dimension, so that, as we have shown, W is pre-homogeneous with respect to V, we also say that V is *simply pre-transitive on* W.

Let W be a pre-principal space with respect to a pre-group V; by the main theorem of AG, we can construct a group G and a transformation-space S, birationally equivalent to V, W and defined over the same field k; then S is also pre-principal with respect to G. Let T be the locus of (u, xu) over k on $S \times S$, x and u being independent generic points of G, S over k; put $t = (u, xu)$; then (TG 1′) implies that $k(x) \subset k(t)$, i. e. that we may write $x = \phi(t)$, where ϕ is a mapping of T into G, defined over k; conversely, if this is so for a transformation-space S with respect to G, S is pre-principal.

The space S will be called a *principal space with respect to* G if, for x, u generic and independent over k on G, S and for $t = (u, xu)$, we have $x = \phi(t)$ where ϕ is an *everywhere defined mapping*, defined over k, of the locus T of t over k into the group G. If at the same time S is homogeneous, it will be called a *principal homogeneous space* with respect to G.

PROPOSITION 3. *Let S be a pre-principal transformation-space with respect to a group G, both being defined over a field k. Then there is a k-open subset P of S which is a principal transformation-space with respect to G; if G and S have the same dimension, P is uniquely determined and is homogeneous.*

Let T and ϕ be defined as above; call F the k-closed subset of T where ϕ is not defined. We first show that, if (a, b) is in F, $(sa, s'b)$ is in F for all s, s' in G. In fact, take x, u generic and independent over $k(s, s')$ on G, S; put $v = xu$, $u_1 = su$, $v_1 = s'v$, $x_1 = s'xs^{-1}$; then we have $v_1 = x_1u_1$, and x_1, u_1 are generic and independent over $k(s, s')$ on G, S, so that (u, v) and (u_1, v_1) are generic points of T over $k(s, s')$, and that $x = \phi(u, v)$, $x_1 = \phi(u_1, v_1)$ by the definition of ϕ; this gives

$$\phi(u, v) = s'^{-1}\phi(su, s'v)s.$$

If (a, b) is in T, it is a specialization of (u, v) over $k(s, s')$, and therefore $(sa, s'b)$ is also in T; then the above relation shows that ϕ is defined at (a, b) if it is defined at $(sa, s'b)$, i.e. that $(a, b) \in F$ implies $(sa, s'b) \in F$.

As (e, u) is a specialization of (x, u) over k, e being the neutral element of G, T contains the diagonal Δ of $S \times S$. As the projection of Δ on either factor of $S \times S$ is everywhere biregular, the projection of the k-closed subset $F \cap \Delta$ of Δ onto S is a k-closed subset F' of S, consisting of the points $a \in S$ such that ϕ is not defined at (a, a). From what we have proved above, it follows that, if $a \in F'$, $sa \in F'$ for all $s \in G$. For the same reason, if a is in $S - F'$, then $\phi(a, sa)$ is defined for all $s \in G$; as (a, sa, s) is then a specialization of (u, xu, x) over k, and $x = \phi(u, xu)$, this shows that $\phi(a, sa) = s$ for all $a \in S - F'$ and all $s \in G$, and therefore $k(a, s) = k(a, sa)$; in particular, if x is generic over $k(a)$ on G, the locus of xa over $k(a)$ has a dimension equal to that of G.

If G is complete, every specialization (a, b) of (u, xu) over k can be extended to a specialization (a, b, s) of (u, xu, x) over k, so that $b = sa$; in other words, every point of T must be of the form (a, sa), with $a \in S$ and $s \in G$; then it follows from what we have proved above that such a point cannot be in F unless a and sa are in F'. Without attempting to decide whether this is still so in the general case, we shall merely show that, if u

is generic over k on S and (u, u') is in F, then u' must be in F'. In fact, suppose that this is not so; take x generic over $k(u, u')$ on G; call X, X' the loci of xu, xu' over $k(u, u')$ on S; by what we have proved above, they have the same dimension, which is that of G. By F-VI$_3$, Th. 11, we have $T \cap (u \times S) = u \times X$; at the same time, since we have shown that (u, xu') is in T, $u \times X'$ is contained in $T \cap (u \times S)$; as X and X' have the same dimension, this implies that $X = X'$. But, as we have shown, since (u, u') is in F, (u, xu') must be in F, and therefore, since F is k-closed, $u \times X'$ must be contained in F; as $X = X'$, this implies that F contains (u, xu), which is generic on T over k, and contradicts the definition of F. One may observe that, if S is pre-homogeneous, this again shows that (a, b) cannot be in F unless a, b are in F'; for, if $a \notin F'$ and x is generic on G over $k(a, b)$, xa has then over $k(a)$ a dimension equal to that of G, and therefore is generic on S over $k(a)$ since in the present case the dimensions of S and G are equal; then if (a, b) is in F, so is (xa, b), and so b must be in F'.

Now replace first S by $S - F'$; as F' is mapped onto itself by all operations of G, $S - F'$ is again a transformation-space with respect to G, defined over k, and satisfies our other assumptions. Writing again S instead of $S - F'$, we see that it is enough to prove our result under the additional assumption that $F' = \emptyset$. If G is complete or S is prehomogeneous, this already implies that S is principal. Otherwise we observe that, since T is also the locus of $(x^{-1}u, u)$ over k and since we have $\phi(x^{-1}u, u) = x$, T and F are mapped onto themselves by the mapping $(u, v) \to (v, u)$ of $S \times S$ onto itself. Call now F'' the "projection" of F on either factor of $S \times S$ (in the sense of F-IV$_3$ and F-VII$_3$, i.e. the closure of the set-theoretic projection); this will be the same, whether we project F onto the first or the second factor, and it is not S by what we have proved above, since F' is empty; it is therefore a k-closed subset of S. By what we have proved, F'' is mapped onto itself by all operations of G. Then $S - F''$ is the principal space whose existence was to be proved.

Finally, assume that the space S from which we first started was prehomogeneous; this means that $T = S \times S$. Let a, b be any two points in $S - F'$; then, if x is generic on G over $k(a, b)$, xa, xb are generic on S over $k(a, b)$, and so there is an isomorphism σ of $k(a, b, xa)$ onto $k(a, b, xb)$ over $k(a, b)$, mapping xa onto xb; then we have $x^\sigma a = xb$, i.e. $b = x^{-1}x^\sigma a$. This shows that $S - F'$ is homogeneous, and also that an open subset of S which is a transformation-space for G cannot contain a point of $S - F'$ without containing $S - F'$. Therefore $S - F'$ is the only open subset of S which is a principal space with respect to G.

If S is a principal homogeneous space, the mapping ϕ of $T = S \times S$ into G which has been defined above will be called the *canonical mapping* of $S \times S$ into G. For any a, b on S and s on G, the relations $b = sa$, $s = \phi(a, b)$ are equivalent; in particular, for any a on S and for x generic over $k(a)$ on G, the mapping $x \rightarrow xa = v$ of G into S has the inverse $v \rightarrow x = \phi(a, v)$; as both are everywhere defined, this is therefore an everywhere biregular mapping of G onto S, defined over $k(a)$. In particular, if there is at least one rational point a over k on S, S is biregularly equivalent to G over k.

4. Let G be a group, V and W two varieties, and F a mapping of $V \times W$ into G, all defined over a field k. We may consider $W \times G$ as a transformation-space with respect to G, the law of composition between them being $(x, (N, y)) \rightarrow (N, xy)$ for any N in W and any x, y in G. We now apply Prop. 1 of no. 1 to the case when we take for S this transformation-space $W \times G$ and for Z the graph of the mapping $N \rightarrow F(M, N)$ of W into G, where M, N are independent generic points of V, W over k. We must then consider the smallest field of definition k' containing k for the mapping $N \rightarrow xF(M, N)$ of W into G, where x is generic over $k(M, N)$ on G. As this mapping is defined over $k(x, M)$, k' is a regular extension of k, contained in $k(x, M)$. Then Prop. 1 shows that we may write $k' = k(u)$, where u is a generic point over k of a transformation-space U with respect to G; as $k(u) \subset k(x, M)$, we may write $u = f(x, M)$, where f is a mapping of $G \times V$ into U, defined over k; moreover, as the mapping $x \rightarrow f(x, M)$ of G into U is no other than the mapping F defined in Prop. 1, we see that f is defined at every point (s, M) of $G \times M$, and that $f(ss', M) = sf(s', M)$; taking $s' = e$, and writing $f(M)$ instead of $f(e, M)$, this gives $f(s, M) = sf(M)$, and in particular $u = xf(M)$. As the mapping $N \rightarrow xF(M, N)$ is defined over $k(u)$, $xF(M, N)$ is rational over $k(u, N)$; similarly, if y is generic over $k(x, M, N)$ on G, the mapping $N \rightarrow yxF(M, N)$ is defined over $k(yu)$, and so $yxF(M, N)$ is rational over $k(yu, N)$. As we can write

$$y = (yxF(M, N))(xF(M, N))^{-1},$$

this shows that y is rational over $k(u, yu, N)$. If N' is generic on W over $k(x, y, M, N)$, y must then also be rational over $k(u, yu, N')$; thus $k(y)$ is contained in $k(u, yu, N)$ and in $k(u, yu, N')$; as these are independent regular extensions of $k(u, yu)$, their intersection is $k(u, yu)$, and so we have $k(y) \subset k(u, yu)$. This means that U is a pre-principal space and may therefore, by Prop. 3, be replaced by a principal space, birationally equivalent to it.

The most interesting case is that in which there are two mappings F_1, F_2 of V, W into G, defined over some overfield K of k, such that $F(M,N)=F_1(M)F_2(N)$ for M, N generic and independent over K on V, W; by the corollary of Th. 7, VA-18, this is always so whenever G is an abelian variety. Take x generic over $K(M,N)$ on G, and put $z=xF_1(M)$; then the mapping $N\rightarrow xF(M,N)=zF_2(N)$ is defined over $K(z)$, so that $k(u)\subset K(z)$. As x, M are generic and independent over $K(N)$ on G, V, u is then generic over $K(N)$ on U, and the dimension of U is that of u over K; the relation $K(u)\subset K(z)$ shows that this is at most the dimension of G. Therefore U is pre-homogeneous and may be taken to be a principal homogeneous space with respect to G. Moreover, we may write $u=\Phi(z)$, where Φ is a mapping of G into U, defined over K. If we substitute yx for x, with y generic over $K(M,N,x)$ on G, z is replaced by yz, and u by yu; this gives $\Phi(yz)=y\Phi(z)$, which may be written as $\Phi(z)=y^{-1}\Phi(yz)$ and thus shows that Φ is everywhere defined. Putting now $a=\Phi(e)$, we see that a is rational over K and that $u=za$, i.e. $f(M)=F_1(M)a$. Put now $g(N)=F_2(N)^{-1}a$; g is a mapping of W into U, defined over K. As we have also $g(N)=F(M,N)^{-1}f(M)$, g is also defined over the field $k(M)$, and therefore also over $k(M')$ if M' is another generic point of V over K; if we take M, M' generic and independent over K on V, $k(M)$ and $k(M')$ are independent regular extensions of k, so that their intersection is k; hence g is defined over k. Thus we have proved the following result:

PROPOSITION 4. *Let G be a group, V and W two varieties and F a mapping of $V\times W$ into G, all defined over k. Assume that there are two mappings F_1, F_2 of V, W into G, defined over some overfield K of k, such that $F(M,N)=F_1(M)F_2(N)$ for M, N generic and independent over K on V, W. Then there is a principal homogeneous space U with respect to G, and two mappings f, g of V, W into U, all defined over k, such that $f(M)=F(M,N)g(N)$, i.e. $F(M,N)=\phi(g(N),f(M))$, where ϕ is the canonical mapping of $U\times U$ into G.*

COROLLARY. *Notations being as in Prop. 4, U, f and g are uniquely determined by G, V, W and F, up to an isomorphism.*

In fact, assume that U', f', g' have similar properties; then we have $xF(M,N)=\phi'(g'(N),xf'(M))$, where ϕ' is the canonical mapping for U'. This shows that the mapping $N\rightarrow xF(M,N)$ is defined over $k(u')$ with $u'=xf'(M)$; thus, if we put $u=xf(M)$ as before, we have $k(u)\subset k(u')$ and may write $u=\psi(u')$, where ψ is a mapping of U' into U, defined over k.

Replacing x by yx, with y generic on G over $k(M,N,x)$, we get $yu=\psi(yu')$; from this we conclude, in the usual manner, that ψ is everywhere defined. Take any point a' on U', and put $a=\psi(a')$; as we have $xa=\psi(xa')$, and as the mappings $x\rightarrow xa$, $x\rightarrow xa'$ are everywhere biregular mappings of G onto U and U', defined over $k(a')$, we see that ψ is an everywhere biregular mapping of U' onto U. Moreover, we have $\psi(u')=x\psi(f'(M))$, and therefore $f=\psi\circ f'$; from this one easily concludes that $g=\psi\circ g'$. This proves our assertion.

5. Let G be a group, defined over a field k. We will now prove that the classes of principal homogeneous spaces with respect to G, for birational equivalence over k, form a set. In fact, let x, y be independent generic points of G over k; let σ be the isomorphism of $\bar{k}(x)$ onto $\bar{k}(yx)$ over $\bar{k}$ which maps x onto yx. Let H be any principal homogeneous space with respect to G, defined over k; let a be an algebraic point over k on H, and put $u=xa$, so that u is generic over k on H. Then $k(u)$ is a regular extension of k contained in $\bar{k}(x)$ and such that $\bar{k}(u)=\bar{k}(x)$; moreover, we have

$$k(y,u)=k(y,u^{\sigma})=k(u,u^{\sigma})$$

since $u^{\sigma}=yu$. Conversely, let $k(u)$ be any such extension of k, and call U the locus of u over k; then we may write $u^{\sigma}=g(y,u)$, where g is a mapping of $G\times U$ into U, defined over k; and one verifies at once that this makes U into a pre-principal pre-homogeneous space with respect to G, and thus determines uniquely a class of birationally equivalent principal homogeneous spaces with respect to G. As every such class is determined by at least one such extension, this shows that these classes form a set.

If G is commutative, one can define canonically a commutative group-structure on the set of classes of principal homogeneous spaces with respect to G. In order to do this, we first observe that, if H is any transformation-space over a commutative group G, then the law $(x,u)\rightarrow x^{-1}u$, for $x\in G$, $u\in H$, defines on H a structure of transformation-space with respect to G; this will be called the *opposite* transformation-space to H and will be denoted by H^{-}; it is a principal homogeneous space if H is such.

PROPOSITION 5. *Let G be a commutative group, defined over a field k. Let H_i, for $1\leqq i\leqq n$, be principal homogeneous spaces with respect to G, defined over k. Then there is a principal homogeneous space H with respect to G, defined over k, and an everywhere defined mapping f of $H_1\times H_2\times\cdots\times H_n$ into H, defined over k, such that*

$$f(s_1a_1,\cdots,s_na_n)=s_1\cdots s_nf(a_1,\cdots,a_n)$$

for all $s_i \in G$ *and* $a_i \in H_i$. *Moreover,* H *and* f *are uniquely determined up to an isomorphism of* H.

Put $V = W = H_1 \times H_2 \times \cdots \times H_n$; call ϕ_i the canonical mapping of $H_i \times H_i$ into G, so that $b_i = sa_i$ is equivalent to $s = \phi_i(a_i, b_i)$ for a_i, b_i in H_i and s in G. Let $u = (u_1, \cdots, u_n)$ and $v = (v_1, \cdots, v_n)$ be two points of V; put

$$F(u, v) = \prod_{i=1}^{n} \phi_i(u_i, v_i),$$

where the right-hand side has a meaning since G is commutative. On each H_i, choose a point a_i, and put $a = (a_1, \cdots, a_n)$. We have

$$\phi_i(u_i, v_i) = \phi_i(a_i, v_i)\phi_i(a_i, u_i)^{-1}$$

for all i, as one verifies at once, and therefore, again because of the commutativity of G:

$$F(u, v) = F(a, v)F(a, u)^{-1}.$$

Thus the assumptions of Prop. 4 are satisfied, so that there is a principal homogeneous space U and two mappings f, g of V into U, all defined over k, such that

$$(1) \qquad f(u) = F(u, v)g(v), \qquad F(u, v) = \phi(g(v), f(u)),$$

where ϕ is the canonical mapping of $U \times U$ into G. Take any point b on V, and take v generic over $k(b)$ on V; as F is defined at (b, v), the relation (1) shows that f is defined at b. Thus f is everywhere defined. As $F(u, u) = e$, the relation (1) gives $g(u) = f(u)$, i.e. $f = g$. If $s_1, \cdots, s_n$ are any elements of G, and we put $s = s_1 \cdots s_n$ and $u' = (s_1u_1, \cdots, s_nu_n)$, we have $F(u', v) = s^{-1}F(u, v)$ and therefore, by (1), $f(u') = s^{-1}f(u)$. If we now put $H = U^-$, i.e. if we take for H the opposite space to U, H and f will have the properties stated in Prop. 5.

Let us now assume that $\bar{H}$ and $\bar{f}$ have similar properties; put $\bar{U} = \bar{H}^-$. Put $\bar{z} = F(u, v)^{-1}\bar{f}(u)$, the multiplication in the right-hand side being that of $\bar{U}$. If the s_i, s and u' have the same meaning as above, we have $\bar{f}(u') = s^{-1}\bar{f}(u)$, so that $\bar{z}$ does not change if one replaces u, v by u', v. Therefore $k(\bar{z})$ is contained both in $k(u, v)$ and in $k(u', v)$. If the s_i have been taken generic and independent over $k(u, v)$ on G, $k(u, v)$ and $k(u', v)$ will be independent regular extensions of $k(v)$; this gives $k(\bar{z}) \subset k(v)$, so that we may write $\bar{z} = \bar{g}(v)$, with $\bar{g}$ defined over k. Then we have $\bar{f}(u) = F(u, v)\bar{g}(v)$; by the corollary of Prop. 4, $\bar{U}$, $\bar{f}$ and $\bar{g}$ must then be the same as U, f and f, respectively, except for an isomorphism of U onto $\bar{U}$. This proves the assertion about unicity in Prop. 5.

In Prop. 5, take $n=2$; call $\mathcal{H}_1$, $\mathcal{H}_2$ the classes of H_1, H_2, and denote by $\mathcal{H}_1+\mathcal{H}_2$ the class of H. This defines on the set of classes of principal homogeneous spaces with respect to G a commutative group-structure. In fact, commutativity is obvious. Call $\mathcal{H}_0$ the class of G, and therefore of all principal homogeneous spaces with respect to G which have a rational point over k. For any principal homogeneous space H with respect to G, the mapping $f(x,u)=xu$ of $G\times H$ into H satisfies the condition of Prop. 5; therefore we have $\mathcal{H}_0+\mathcal{H}=\mathcal{H}$ for all classes $\mathcal{H}$. If ϕ is the canonical mapping of $H\times H$ into G, then ϕ, considered as a mapping of $H\times H^-$ into G, satisfies the condition of Prop. 5; therefore, if $\mathcal{H}^-$ is the class of H^-, we have $\mathcal{H}+\mathcal{H}^-=\mathcal{H}_0$. Finally, let H_1, H_2, H_3 be three principal homogeneous spaces with respect to G; apply Prop. 5 successively to the following spaces: (a) to H_1, H_2, obtaining a space H_{12} and a mapping f_{12}; (b) to H_{12}, H_3, obtaining H', f'; (c) to H_2, H_3, obtaining H_{23}, f_{23}; (d) to H_1, H_{23}, obtaining H'', f''; (e) to H_1, H_2, H_3, obtaining H, f. Then the two mappings

$$f'(f_{12}(u_1,u_2),u_3),\qquad f''(u_1,f_{23}(u_2,u_3))$$

of $H_1\times H_2\times H_3$ into H', H'' satisfy the same condition as the mapping f. By the unicity assertion of Prop. 5, this shows that H', H'' are isomorphic to H. This means that the addition $\mathcal{H}_1+\mathcal{H}_2$ is associative.

One proves quite similarly, by induction on n, that if $\mathcal{H}$ and $\mathcal{H}_i$ are the classes of the spaces H, H_i in Prop. 5, then $\mathcal{H}=\sum_i \mathcal{H}_i$. In fact, let H', f' be the space and the mapping obtained by applying Prop. 5 to $H_1,\cdots,H_{n-1}$, so that $\mathcal{H}'=\mathcal{H}_1+\cdots+\mathcal{H}_{n-1}$ by the induction assumption; and let H'', f'' be the space and the mapping obtained by applying Prop. 5 to H', H_n, so that $\mathcal{H}''=\mathcal{H}'+\mathcal{H}_n$ by definition. Then the mapping

$$(u_1,\cdots,u_n)\rightarrow f''(f'(u_1,\cdots,u_{n-1}),u_n)$$

of $H_1\times\cdots\times H_n$ into H'' has the properties stated for f in Prop. 5, so that, by the unicity assertion in Prop. 5, H'' is isomorphic to H.

From this one deduces that every element $\mathcal{H}$ of the group we have just described is of finite order. In fact, take on a space H of class $\mathcal{H}$ any positive cycle $\sum_{i=1}^{n} a_i$ of dimension 0, rational over k. Call H_n any space of class $n\mathcal{H}$; then there is a mapping $f(u_1,\cdots,u_n)$ of the product of n factors equal to H into H_n with the properties stated in Prop. 5. From the unicity assertion in Prop. 5, it follows that any permutation of the u_i will change f

into sf, with $s \in G$; as f is everywhere defined, we see that $s = e$ by taking $u_1 = \cdots = u_n$; therefore f is a symmetric function, so that $f(a_1, \cdots, a_n)$ is rational by the main theorem on symmetric functions (VA-7, Th. 1). So H_n has a rational point over k, and is therefore isomorphic to G.

Now, $\mathfrak{A}$ being as before, put $H_0 = G$ and take, for each integer $n \neq 0$, a space H_n of class $n\mathfrak{A}$ so that all the H_n are disjoint. On the set $\mathfrak{G} = \bigcup_n H_n$ (which is of course not an algebraic variety), we will define a commutative group-law f (in the sense of group-theory, not of algebraic geometry) such that $H_0 = G$ will be a subgroup of $\mathfrak{G}$ and that f induces on $H_m \times H_n$, for all m, n, a mapping $f_{m,n}$ of $H_m \times H_n$ into H_{m+n} satisfying the conditions in Prop. 5. As there is such a mapping $f_{m,n}$ for each m, n, and as it is uniquely determined up to an automorphism of H_{m+n} (i. e. up to left-multiplication by a rational point of G), we merely have to choose the $f_{m,n}$ so that the mapping f of $\mathfrak{G} \times \mathfrak{G}$ into $\mathfrak{G}$ which coincides with $f_{m,n}$ on $H_m \times H_n$ for all m, n satisfies the axioms for groups; we do this as follows. For any n, we take $f_{0,n}(x, u) = xu$ for $x \in G$, $u \in H_n$. We choose $f_{-1,1}$ and, for all $n > 0$, $f_{n,1}$ and $f_{-n,-1}$ so as to satisfy the conditions in Prop. 5. Now, for elements $u_1, \cdots, u_{n+1}$ of H_1 in any number, we define $u_1 \cdots u_{n+1}$ inductively as being equal to u_1 for $n = 0$ and to $f_{n,1}(u_1 \cdots u_n, u_{n+1})$ for $n \geqq 1$; similarly, for elements $v_1, \cdots, v_{n+1}$ of H_{-1}, we define $v_1 \cdots v_{n+1}$ as equal to v_1 for $n = 0$ and to $f_{-n,-1}(v_1 \cdots v_n, v_{n+1})$ for $n \geqq 1$. It is then easily seen that, whenever m, n are both > 0, there is one and only one way of choosing $f_{m,n}$ so that it satisfies the condition

$$f_{m,n}(u_1 \cdots u_m, u_{m+1} \cdots u_{m+n}) = u_1 \cdots u_{m+n}$$

when the u_i are in H_1; we determine $f_{-m,-n}$ similarly, using H_{-1} instead of H_1. Finally, for $m \geqq n > 0$, we choose $f_{m,-n}$ and $f_{-m,n}$ so as to satisfy the conditions

$$f_{m,-n}(u_1 \cdots u_m, v_1 \cdots v_n) = \prod_{i=1}^{n} f_{-1,1}(v_i, u_i) \cdot u_{n+1} \cdots u_m$$

$$f_{-m,n}(v_1 \cdots v_m, u_1 \cdots u_n) = \prod_{i=1}^{n} f_{-1,1}(v_i, u_i) \cdot v_{n+1} \cdots v_m$$

respectively, the u_i being any elements of H_1 and the v_i any elements of H_{-1}. It is then a trivial matter to verify that these choices of the $f_{m,n}$ satisfy all the requirements for a commutative group-law on $\mathfrak{G}$.

The points on the H_n which are rational over k form a subgroup $\mathfrak{g}$ of $\mathfrak{G}$. As we have shown that there are such points for some $n \neq 0$, there is a smallest $n > 0$ for which there is such a point $a \in H_n$; this n is the order of $\mathfrak{A}$ in the group of classes of principal homogeneous spaces with respect to G. Then $\mathfrak{g}$ is the direct product of the group $\mathfrak{g}_0 = \mathfrak{g} \cap G$ of rational points over k on G and of the infinite cyclic group γ generated by a. The quotient-group $\mathfrak{G}/\gamma$ may be described as an algebraic group consisting of n components respectively isomorphic to $H_0 = G, H_1, \cdots, H_{n-1}$.

6. PROPOSITION 6. *Let A be an abelian variety and H a principal homogeneous space with respect to A, both being defined over a field k. Let $V_1, \cdots, V_n$ be varieties, and F a mapping of $V_1 \times \cdots \times V_n$ into H, all these being defined over k. Then there is for each i a principal homogeneous space H_i with respect to A and a mapping F_i of V_i into H_i, H_i and F_i being defined over k, and there is a mapping f of $H_1 \times \cdots \times H_n$ into H with the properties stated in Prop. 5, such that, for $(M_1, \cdots, M_n)$ generic over k on $V_1 \times \cdots \times V_n$, we have*

$$F(M_1, \cdots, M_n) = f(F_1(M_1), \cdots, F_n(M_n)).$$

Moreover, all these are uniquely determined up to isomorphisms.

For $n = 1$, there is nothing to prove. If the assertion is proved for a product of two factors, then this can be applied to the product $V_1 \times (V_2 \times \cdots \times V_n)$ of V_1 and $V_2 \times \cdots \times V_n$, so that the general case follows by induction on n. Thus it is enough to treat the case of two factors V, W and of a mapping F of $V \times W$ into H. Call ϕ the canonical mapping of $H \times H$ into A; let (M, N) and (M', N') be two independent generic points of $V \times W$ over k; and put

$$x = \phi(F(M, N'), F(M', N')), \qquad y = \phi(F(M, N), F(M, N')).$$

so that we have

$$xy = \phi(F(M, N), F(M', N')).$$

As the mapping $((M, M'), N') \rightarrow x$ of $(V \times V) \times W$ into A has the constant value e on the variety $(M, M) \times W$, Th. 7 of VA-18 shows that x is rational over $k(M, M')$; for a similar reason, y must be rational over $k(N, N')$; in other words, there are mappings Φ, Ψ of $V \times V$, $W \times W$ into A, both defined over k, such that $x = \Phi(M, M')$ and $y = \Psi(N, N')$. By the corollary of Th. 7 of VA-18, Φ and Ψ satisfy the assumptions of Prop. 4, so that there are two principal homogeneous spaces U_1, U_2 with respect to A, mappings F_1, G_1 of V into U_1 and mappings F_2, G_2 of W into U_2, all defined over k, such that $F_1(M) = xG_1(M')$, $F_2(N) = yG_2(N')$; moreover, as $\Phi(M, M)$, $\Psi(N, N)$ are defined and equal to e, we have $G_1 = F_1$, $G_2 = F_2$. Now call H_1, H_2 the spaces respectively opposite to U_1, U_2, and apply Prop. 5 to H_1, H_2: let $\bar{H}$ be the principal homogeneous space and $\bar{f}$ the mapping of $H_1 \times H_2$ into $\bar{H}$ with the properties stated in that proposition. Put $\bar{F}(M,N)=\bar{f}(F_1(M),F_2(N))$. As H_1, H_2 are opposite to U_1, U_2, we have

$$F_1(M')=xF_1(M), \qquad F_2(N')=yF_2(N),$$

multiplication in the right-hand sides being understood in the sense of H_1, H_2. By the definition of $\bar{f}$, we have then:

$$\bar{F}(M',N')=(xy)\bar{F}(M,N),$$

while, as we have seen above, the same relation holds if F is substituted for $\bar{F}$. But then, as the corollary of Th. 7, VA-18, shows that the mapping

$$((M',N'),(M,N))\rightarrow xy$$

of $(V\times W)\times(V\times W)$ into A satisfies the condition of Prop. 4, the corollary of Prop. 4 shows that $\bar{H}$, $\bar{F}$ must be the same as H, F except for an isomorphism of H onto $\bar{H}$. Then, replacing $\bar{f}$ by a mapping f of $H_1\times H_2$ into H by means of that isomorphism, we have the spaces H_1, H_2 and the mappings F_1, F_2, f whose existence was asserted in our proposition.

As to unicity, assume that there are spaces $H_1{}^*$, $H_2{}^*$ and mappings $F_1{}^*$, $F_2{}^*$, f^* with the same properties. Then, x being defined as above, or equivalently by $F(M',N')=xF(M,N')$, we have

$$f^*(F_1{}^*(M'),F_2{}^*(N'))=xf^*(F_1{}^*(M),F_2{}^*(N'))=f^*(xF_1{}^*(M),F_2{}^*(N'))$$

and therefore $F_1{}^*(M')=xF_1{}^*(M)$ since the mapping $u\rightarrow f^*(u,v)$ of $H_1{}^*$ into H is, as easily seen, an everywhere biregular mapping of $H_1{}^*$ onto H. But then the corollary of Prop. 4, applied to the mapping $(M',M)\rightarrow x$ of $V\times V$ into A, shows that $H_1{}^*$, $F_1{}^*$ are the same as H_1, F_1 except for an isomorphism. The same argument applied to y instead of x shows that H_2, F_2 are uniquely determined up to an isomorphism. Then Prop. 5 shows that f is uniquely determined. This completes the proof.

7. The foregoing results will now be applied to the theory of Jacobian varieties. As in VA-35, we consider a complete non-singular curve Γ of genus $g>0$, defined over a field k. If $\mathfrak{a}$ is any divisor on Γ, Prop. 6 of the Appendix of AG shows that there is a smallest field containing k over which $\mathfrak{a}$ is rational; this field will be denoted by $k(\mathfrak{a})$. In particular, if $M_1,\cdots,M_g$ are independent generic points of Γ over k and if we put $\mathfrak{m}=\sum_i M_i$, then, by VA-4, Lemma 1, $k(\mathfrak{m})$ is the field $k(M_1,\cdots,M_g)_s$ of symmetric functions of $M_1,\cdots,M_g$ defined over k, i.e. the subfield of $k(M_1,\cdots,M_g)$ consisting of those elements which are invariant under all permutations of

$M_1, \cdots, M_g$; such a divisor $\mathfrak{m}$ will be called generic over k. As $k(\mathfrak{m})$ is a regular extension of k, we may write it as $k(u)$, where u is a generic point of a variety W over k, and we may write $u = F(M_1, \cdots, M_g)$, with F defined over k; as F is symmetric in the M_i, this may also be written as $u = F(\mathfrak{m})$.

Now let the N_i, P_i, for $1 \leqq i \leqq g$, be $2g$ independent generic points of Γ over $k(\mathfrak{m})$; and put:

$$x = (N_1, \cdots, N_g, P_1, \cdots, P_g),$$

this being a generic point over k of the product $V = \Gamma \times \cdots \times \Gamma$ of $2g$ factors equal to Γ. By VA-35, Lemma 11, there is a positive divisor $\mathfrak{m}'$ on Γ linearly equivalent to $\mathfrak{m} + \sum_{i=1}^{g} N_i - \sum_{i=1}^{g} P_i$, and it is uniquely determined and such that $k(x, \mathfrak{m}) = k(x, \mathfrak{m}')$; this implies that it is generic over $k(x)$. Then, if we write $u' = F(\mathfrak{m}')$, we have $k(x, u) = k(x, u')$; we may thus write $u' = g(x, u)$, where g is a mapping of $V \times W$ into W which satisfies (TG 1).

We now show that this mapping satisfies the condition (TG 2′) of AG, no. 3, Prop. 2, so that this proposition may be applied to it. In fact, let y be a generic point of V over $k(x, u)$; we may write

$$y = (Q_1, \cdots, Q_g, R_1, \cdots, R_g).$$

Then the point $u'' = g(y, u')$ will be determined by $u'' = F(\mathfrak{m}'')$, where $\mathfrak{m}''$ is the positive divisor linearly equivalent to $\mathfrak{m}' + \sum_i Q_i - \sum_i R_i$. Applying again VA-35, Lemma 11, we see that there is a positive divisor $\sum_i S_i$ linearly equivalent to $\sum_i N_i - \sum_i P_i + \sum_i Q_i$, and that it is generic over $k(y, u)$ and rational over $k(x, y)$. But then $\mathfrak{m}''$ is linearly equivalent to $\mathfrak{m} + \sum_i S_i - \sum_i R_i$, which shows that, if we put

$$z = (S_1, \cdots, S_g, R_1, \cdots, R_g)$$

z is generic on V over $k(u)$ and that we have $u'' = g(z, u)$. This shows that g satisfies (TG 2′). Applying Prop. 2 of AG, no. 3 and the main theorem of AG, we see that there is a group J, a normal law $\bar{g}$ between J and W, and a mapping ϕ of V into J such that $g(x, u) = \bar{g}(\phi(x), u)$ for x, u generic and independent over k on V, W.

Put now $K = k(P_1, \cdots, P_g)$, $\mathfrak{n} = \sum_i N_i$, and

$$w = (S_1, \cdots, S_g, Q_1, \cdots, Q_g).$$

Since the P_i, Q_i, N_i are generic and independent over k, Lemma 11 of VA-35 shows that w is generic over K on V. At the same time, the linear equivalence by which the S_i were defined shows at once that $g(x,u)=g(w,u)$ for u generic over $k(x,w)$, and therefore $\bar{g}(\phi(x),u)=\bar{g}(\phi(w),u)$. Since J, by definition, operates faithfully on W, this implies $\phi(x)=\phi(w)$; as w is generic over K on V, this shows that $\phi(x)$ is generic over K on J. It will now be shown that $K(\phi(x))=K(\mathfrak{n})$. In fact, if u and $u'=g(x,u)$ are as before, u' is rational over $K(\mathfrak{m},\mathfrak{n})=K(u,\mathfrak{n})$ since the divisor $\mathfrak{m}'$ is so by Lemma 11 of VA-35; therefore the mapping $u\rightarrow u'$ is defined over $K(\mathfrak{n})$, so that $K(\phi(x))\subset K(\mathfrak{n})$. Put now $K'=K(\phi(x))$, so that u' is rational over $K'(u)$; then $\mathfrak{m}$ and $\mathfrak{m}'$ are both rational over $K'(u)$. But Lemma 11 of VA-35 shows that $\mathfrak{n}$ is rational over $K(\mathfrak{m},\mathfrak{m}')$ and therefore over $K'(u)$. If u_1 is a generic point of W over $K'(u)$, $\mathfrak{n}$ will also be rational over $K'(u_1)$; as $K'(u)$, $K'(u_1)$ are independent regular extensions of K', this implies that $\mathfrak{n}$ is rational over K'.

But now a comparison with the construction of the jacobian variety given in VA-36 shows that the latter coincides with our J over a suitably extended groundfield; more precisely, substituting K for k, $\sum_i P_i$ for $\mathfrak{a}$, $\mathfrak{n}$ for $\mathfrak{m}$ and $\phi(x)$ for z in the treatment given in VA-36, we get the same law of composition for the field $K(\phi(x))$ as has been defined above. Alternatively, one may also reason as follows. Let J_1 be the jacobian variety as defined in VA; let ϕ_1 be the "canonical mapping" of Γ into J_1, also according to the definition of VA-37 (which will soon be replaced by a more appropriate one); let K_1 be an overfield of the field K defined above, over which J_1 and ϕ_1 are defined; take $\mathfrak{n}$ generic over K_1; put $t=\phi(x)$, x being as above, and $z=S[\phi_1(\mathfrak{n})]$. As we have then $K_1(x)=K_1(z)=K_1(\mathfrak{n})$, the mapping $x\rightarrow z$ defines a birational correspondence between J and J_1, defined over K_1. If we write it as $z=f(x)$, f is everywhere defined by VA-15, Th. 6; and this, by the results at the beginning of VA-19, must then be of the form $f(x)=f_0(x)+a$, where $a=f(e)$ and where f_0 is a homomorphism, so that (using the additive notation on J_1 and the multiplicative notation on J) we have $f_0(xx')=f_0(x)+f_0(x')$. But then f_0 is again a birational correspondence, and, if g is the inverse mapping to f_0, we have $g(z+z')=g(z)g(z')$ for z, z' generic and independent over K_1 on J_1. This can be written as $g(z)=g(z+z')g(z')^{-1}$; if then z_1 is any point of J_1, and we take z' generic on J_1 over $K_1(z_1)$, this shows that g is defined at z_1. As f_0, g are everywhere defined, they determine an isomorphism between J and J_1.

One could also, without making use of the results of VA, verify directly (for instance by making use of the criterion for the completeness of a group given by VA-33, Th. 16) that the group J we have constructed here is complete and is therefore an abelian variety. Then the results we have proved above, combined with the corollary of Th. 7, VA-18, show at once that J has the properties stated in VA-36, Th. 18; since the whole theory of the jacobian variety depends upon nothing else, and these properties (as proved in VA-37) are characteristic of the jacobian variety, this would suffice for a complete treatment.

From this discussion, we conclude that J is an abelian variety. Now apply Prop. 6 to the mapping ϕ of V into J; this defines $2g$ mappings of Γ into principal homogeneous spaces with respect to J, all defined over k. As ϕ is symmetric in the N_i and also in the P_i, the unicity assertion in Prop. 6 shows at once that the first g mappings must coincide, and that the last g mappings must coincide; call F, F' these mappings, and H, H' the spaces into which they map Γ. Now, notations being the same as above in no. 6, put

$$x' = (P_1, \cdots, P_g, N_1, \cdots, N_g).$$

Then we have, always with the same notations as before, $u = g(x', u')$, and therefore $\phi(x') = \phi(x)^{-1}$. This, combined with the unicity assertion in Prop. 6, shows at once that H' is the opposite space to H while F' must be the same as F.

We now embed J and H, in the manner explained at the end of no. 5, into a commutative group $\mathfrak{G}$ consisting of principal homogeneous spaces H_n with respect to J, all defined over k, with $H_0 = J$, $H_1 = H$, in such a way that $\mathfrak{G}/J$ is an infinite cyclic group, that the H_n are the cosets of J in $\mathfrak{G}$ and that the group-law in $\mathfrak{G}$ induces on $H_m \times H_n$, for all m, n, a mapping of $H_m \times H_n$ into H_{m+n} defined over k and satisfying the conditions in Prop. 5. At the same time, we change from the multiplicative to the additive notation, not only in J but also in $\mathfrak{G}$. With this notation, we have, if x, $\phi(x)$ and F have the same meaning as before,

$$\phi(x) = \sum_i F(N_i) - \sum_i F(P_i).$$

Let us now extend the mapping F into a homomorphism of the group of divisors on Γ into $\mathfrak{G}$, by putting $F(\mathfrak{a}) = \sum_i n_i F(A_i)$ for any divisor $\mathfrak{a} = \sum_i n_i A_i$, so that $F(\mathfrak{a}) \in H_n$ if n is the degree of $\mathfrak{a}$; in particular, $F(\mathfrak{a})$ is in J if and only if $\mathfrak{a}$ is of degree 0. If a is any point of H and M is a generic point of Γ over $k(a)$, the mapping $M \rightarrow F(M) - a$ of Γ into J, which is defined over $k(a)$, is a "canonical mapping" in the sense of VA-37; naturally it is only defined up to an additive constant; and, by the unicity assertion in Prop. 6, no such mapping can be defined over k unless H is isomorphic to G, i. e. unless H has a rational point over k. From Th. 19 of VA-38, one deduces immediately that a divisor $\mathfrak{a}$ on Γ is linearly equivalent to 0 if and only if $F(\mathfrak{a})=0$.

In other words, *the homomorphism* $\mathfrak{a}\rightarrow F(\mathfrak{a})$ *determines an isomorphism of the group of all divisor-classes* (of any degree) *on* Γ *onto the group* $\mathfrak{G}$. From the foregoing results, one concludes easily that these properties are characteristic for $\mathfrak{G}$ and F, up to isomorphisms on J and its cosets H_n in $\mathfrak{G}$. One may call $\mathfrak{G}$ the *Jacobian group of* Γ, and F the *canonical mapping* of Γ, and of the group of divisors on Γ, into the Jacobian group. In substance, the construction of the varieties H_n has already been given by Chow (in [1]) by a method belonging to projective geometry.

THE UNIVERSITY OF CHICAGO.

BIBLIOGRAPHY.

[1] W. L. Chow, "The Jacobian variety of an algebraic curve," *American Journal of Mathematics*, vol. 76 (1954), pp. 453-476.

[2] S. Nakano, "Note on group varieties," *Memoirs of the College of Science, University of Kyoto*, Series A, vol. 27 (1952), Math. no. 1, pp. 55-66.

[3] A. Weil, *Variétés abéliennes et courbes algébriques*, Paris, Hermann et Cie, 1948.

[4] ———, "On algebraic groups of transformations," *American Journal of Mathematics*, vol. 77 (1955), pp. 355-391.

A PROOF OF A THEOREM OF MEYER ON INDEFINITE TERNARY QUADRATIC FORMS.*

By Burton W. Jones and Donald Marsh.

1. Introduction.[1] We shall be considering in this paper primitive indefinite ternary quadratic forms $f = \sum_{i,j=1}^{3} a_{ij}x_ix_j$ where the a_{ij} are integers with g. c. d. 1. We call J the g. c. d. of the two-by-two minor determinants of (a_{ij}) and $\det(a_{ij}) = J^2K$ defines an invariant K. (Classically these invariants were denoted by Ω and Δ). The purpose of this paper is to give a proof of the following theorem of A. Meyer [4] which is very fundamental in the theory of quadratic forms.

Theorem A. *If two primitive indefinite ternary quadratic forms are in the same genus and if their invariants J and K have g. c. d* 1 *or* 2 *and neither is divisible by* 4, *then the two forms are equivalent.*

Meyer gave what he called a proof of this theorem but which contained various obscurities which neither L. E. Dickson [1] nor the authors have been able to resolve. Bachmann did little to clarify these obscurities. L. E. Dickson gave a proof of Meyer's theorem (cf. [1], pp. 35-60) which was very involved in detail and had a few obscurities of its own, though it was an improvement over Meyer's discussion. Oral communication with Martin Eichler and Martin Kneser indicates that this theorem may also be proved using the deep theory of "spinor genera" developed by the former [2].

In this paper, by free use of matrices and the modern theory of quadratic forms including the Hasse symbol, we present a proof which involves a minimum of detailed computation and which we hope is largely free from obscurities.

The outline of the proof is as follows: Theorems 1 and 2 show that two ternary forms are equivalent if they have the same invariants J and K, if they are both properly or both improperly primitive and if they represent a binary form with certain properties. Theorem 3 gives two conditions under which a ternary form f represent all binary forms of determinant JM where

---

* Received February 11, 1955.

[1] This was written under the support of the Office of Naval Research.

M is a prime or double a prime according as the reciprocal form of f is properly or improperly primitive. Theorems 4, 5, 6, and 7 establish the satisfaction of the first of these conditions and Theorem 8 the second. The remainder of the proof is concerned with showing that any two forms of the same genus represent a binary form of determinant JM where M has the required properties.

THEOREM 1. *Let $B = (b_{rs})$ be the matrix of a properly or improperly primitive binary form, h, of determinant JM and let K be an integer such that $g = (M, K)$ is square-free and prime to both M/g and K/g; then there exists a ternary form f with invariants J and K which represents h primitively if and only if there are integers c_1 and c_2 such that*

$$(1) \qquad c_1^2 \equiv -b_{11}K, \qquad c_2^2 \equiv -b_{22}K, \qquad c_1c_2 \equiv +b_{12}K, \qquad (\operatorname{mod} M).$$

Proof. Let A be the matrix of the form f. The f represents h primitively if and only if there is a 3 by 2 integral matrix P whose two-rowed minor determinants have g. c. d. 1 and for which $P^TAP = B$. Then there is an integral column matrix Q such that $U = (P, Q)$ is a unimodular matrix and

$$U^TAU = E = (b_{ij}), \qquad i, j = 1, 2, 3.$$

That is, f represents h if and only if it is equivalent to a form of matrix E whose leading two-by-two minor is B.

First suppose f with invariants J and K represent h, where JM is the determinant of h and (b_{ij}) is the matrix of $h, i, j = 1, 2$. Then write $\operatorname{adj} E = (B_{ij})$ where $B_{33} = JM$ and $B_{ij} \equiv 0 (\operatorname{mod} J)$ for all i, j. Now $\operatorname{adj}(\operatorname{adj} E) = J^2KE$. Hence

$$(2) \qquad B_{22}B_{33} - B^2_{23} = J^2Kb_{11},$$

$$(3) \qquad B_{11}B_{33} - B^2_{13} = J^2Kb_{22},$$

$$(4) \qquad B_{12}B_{33} - B_{13}B_{23} = -J^2Kb_{12},$$

which show that we may take $c_1 = B_{23}/J$, $c_2 = B_{13}/J$ and have the congruences (1).

Conversely, if the congruences (1) hold equations (2), (3), (4) with $c_1 = B_{23}/J$, $c_2 = B_{13}/J$ are solvable for B_{22}, B_{11} and B_{12}. Thus the matrix (B_{ij}) exists. Suppose the elements B_{ij}/J had a common prime factor q. Then q^2 would divide Kb_{11}, Kb_{22}, Kb_{12} and since $(b_{11}, b_{22}, b_{12}) = 1$, q^2 would divide K. But $B_{33} = JM$ and hence $q \mid M$. However q would not be prime to (M, K) unless q^2 divides M which is impossible since (M, K) is square-free. Thus 1 is the g. c. d. of B_{ij}/J. A little calculation shows that J^4K^2 is the determinant of (B_{ij}). Then $\operatorname{adj}(B_{ij})/J^2K$ is equal to a matrix E whose leading two-by-two minor is B and which has the invariants J and K.

Let us henceforth take $(J,K) \leqq 2$, $J \not\equiv 0 \not\equiv K \pmod 4$, and let $w = 1$ or 2 as f is properly or improperly primitive. Notice that if $w = 2$, $K \equiv 2 \pmod 4$ and J is odd.

THEOREM 2. *Two ternary quadratic forms of the same invariants, J, K and w, are equivalent if they both represent primitively the same primitive binary form of determinant JM in which M is an odd prime or twice an odd prime such that $(M,K) = g$ and $(g, K/g) = 1$.*

Proof. For M a prime or twice a prime and $(g, K/g) = 1$, the preceding theorem holds and the congruential conditions (1) have only the pair of solutions c_1, c_2 and the corresponding $-c_1$, $-c_2$, modulo M. The corresponding ternary forms obtained are related by the transformation $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, whereas the form obtained by using $c_1 + tM$, $c_2 + uM$ with t, u integers in place of c_1, c_2 is related to the latter by means of the unimodular transformation

$$\begin{pmatrix} 1 & 0 & -u \\ 0 & 1 & -t \\ 0 & 0 & 1 \end{pmatrix}.$$

We now seek to determine conditions under which the hypothesis of Theorem 2 holds.

2. THEOREM 3. *If a ternary form f represents a primitive binary form $-h$ of determinant JM where M is a prime or double a prime according as the reciprocal form of f is properly or improperly primitive, and if h' is another binary form of determinant JM, then f represents $-h'$ if there are two primes p and q not dividing $2M\,|f|$ such that*

(i) *If T, U is the least solution of $x^2 - wpqKy^2 = 1$, then $T \not\equiv \pm 1 \pmod{wpq}$.*

(ii) *The equations $h = wpq$, $h' = wpq$ are both solvable.*

Proof. Consider h and h' as having been transformed so that their leading coefficients are wpq and let $B = (b_{rs})$ be the matrix of h with $b_{11} = wpq$. Since p, q do not divide M, b_{12} is prime to b_{11} and is one of the four roots of $x^2 \equiv -JM \pmod{wpq}$. If we show that some unimodular transformation of f leaves b_{11} and $|B|$ fixed, but replaces b_{12} by $b_{12}' \not\equiv \pm b_{12}$

(mod wpq), then f represents $-h'$ since the "b_{12}" for h' is one of b_{12}, b_{12}', $-b_{12}$, $-b_{12}'$.

Write the matrix of f in the form (a_{rs}) where $a_{11}=-b_{11}$, consider the reciprocal form and denote

$$\begin{pmatrix} A_{22} & A_{23} \\ A_{23} & A_{33} \end{pmatrix}$$

by R. The determinant of R is $-b_{11}K$. If P is an automorph of R, then the transformation $\begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix}$ on F is equivalent to the transformation $\begin{pmatrix} 1 & 0 \\ 0 & P^{IT} \end{pmatrix}$ on f (leaving b_{11} and $b_{11}K$ invariant). P^I is of the form

$$\begin{pmatrix} t-A_{23}u & -A_{33}u \\ A_{22}u & t+A_{23}u \end{pmatrix}$$

where t, u is a solution of $x^2-wpqKy^2=1$. Now

$$b_{12}'=(a_{12},a_{13})\begin{pmatrix} t-A_{23}u \\ -A_{33}u \end{pmatrix}=a_{12}t-u(a_{12}A_{23}+a_{13}A_{33})$$
$$\equiv a_{12}t-u(a_{11}A_{31}+a_{12}A_{32}+a_{13}A_{33})\equiv a_{12}t \pmod{wpq}.$$

Thus $b_{12}'\equiv b_{12}t \pmod{wpq}$ and, if $t\not\equiv\pm 1 \pmod{wpq}$, then $b_{12}'\not\equiv\pm b_{12} \pmod{wpq}$.

We now find the conditions which must be imposed on p and q if (i) and (ii) of Theorem 3 are to hold. First in the next four theorems we show that if the product pq satisfies certain conditions (mod 8) and (mod K), and if for some fixed factor d_0 of K, p satisfies the condition $(-d_0\,|\,p)=-1$, then

$$x^2-aKy^2=1, \qquad a=wpq \tag{5}$$

has a minimum solution T, U with $T\not\equiv\pm 1 \pmod{wpq}$.

THEOREM 4. *If $K=2^eG^2H$ where GH is odd, H square-free, $(a,K)\leqq 2$ and a is odd or double an odd according as $w=1$ or $w=2$, and $T^2-aKU^2=1$ with $T\equiv\pm 1 \pmod a$, then for some factorization dd' of H with d, d' relatively prime and positive, the following equation has a solution where cc' is the highest power of 2 in 2^eU^2 and not both c, c' are divisible by 4:*

$$dcx^2-d'c'ay^2=\pm 2. \tag{6}$$

Proof. Equation (5) implies $(T\mp 1)(T\pm 1)=2^eaG^2HU^2$ and $T\equiv\pm 1 \pmod a$ implies

$$T\mp 1=ad'c'r_1^2, \qquad T\pm 1=dcr_2^2, \qquad (r_1,r_2)=1, \tag{7}$$

since $T \mp 1$ and $T \pm 1$ have at most a factor 2 in common, where r_1 and r_2 are odd, $r_1^2 r_2^2 cc' = 2^e G^2 U^2$ and $dd' = H$. Subtracting the first equation of (7) from the second, we have

$$(8) \qquad (T \pm 1) - (T \mp 1) = \pm 2 = dcr_2^2 - ad'c'r_1^2,$$

which completes the proof.

Since equations (6) differ according to the parity of e, we specialize these results to find

(1) If $e = 0$, c and c' may both be 1 or one is 2 and the other an odd power of 2 greater than the first; in the former case the equation (6) becomes

$$(6_1) \qquad dx^2 - d'ay^2 = \pm 2.$$

If c and c' are both even, we have

$$(6_2) \qquad dx^2 - d'ay^2 = \pm 1$$

where x or y is even.

(2) If $e = 1$, one of c, c' is 2 and the other is an even power of 2. Then (8) becomes $\pm 1 = dr_2^2 - 2ad'(\frac{1}{2}\sqrt{c'})^2 r_1^2$ or $\pm 2 = d(\sqrt{c})^2 r_2^2 - 2ad'r_1^2$; hence we have one of

$$(6_3) \qquad dx^2 - 2ad'y^2 = \pm 1, \qquad\qquad (6_4) \qquad dx^2 - 2ad'y^2 = \pm 2.$$

We show that by proper choice of a modulo the primes dividing G but not H we may exclude from consideration equations (6_2) and (6_3) with $d = 1$, and $\pm 1 = 1$. Note that $d = 1$ implies $d' = H$.

THEOREM 5. *If, according as e is* 0 *or* 1, *a or $2a$ is so chosen that* $(-aH \,|\, p_i) = -1$ *or* $(-2aH \,|\, p_i) = -1$ *for all primes p_i dividing G but not H, then, if T, U is the least solution of* (5), *it follows that if* $T \equiv \pm 1$ $(\text{mod}\, a)$ *then* (6) *does not reduce to either of the following in the respective cases*:

$$(9) \qquad x^2 - aHy^2 = 1, \qquad x^2 - 2aHy^2 = 1.$$

Proof. Suppose for the least solution T, U equations (7) with $d = 1$, $d' = H$ define r_2, r_1 a solution of (6_2) or (6_3) with $d = 1$ and $\pm 1 = 1$, that is, a solution of one of the equations (9). First we take the case $r_1 = Gr$ and show that T, U is not the least solution of (5). For $e = 0$, $x_0 = r_2(\frac{1}{2}c)^{\frac{1}{2}}$, $y_0 = r_1(\frac{1}{2}c')^{\frac{1}{2}}$ in (9) gives $x_0^2 - aG^2H(y_0/G)^2 = 1$; $x_0, y_0/G$ is a solution of (5) in which $T = cr_2^2 - 1$ and $x_0 = r_2\sqrt{(\frac{1}{2}c)} < T$ which yields our contradiction.

If $e = 1$, $r_1 = Gr$ implies that $x_0 = r_2$, $y_0 = (\frac{1}{2}\sqrt{c'})r$ is a solution of (9) with $x_0 < T$.

Second, suppose that $r_1 \not\equiv 0 \bmod G)$, thus $r_1 r_2 \equiv 0 \pmod{G}$ implies that r_2 contains a prime factor of G and (6_3) and (6_2) with $d = 1$, $d' = H$ implies that the prime factors of G which r_2 contains cannot divide H. For (9) to have no solutions with x divisible by a prime factor of G not dividing H it is therefore sufficient that $(-aH \mid p_i) = -1$ or $(-2aH \mid p_i) = -1$ in the respective cases for all primes p_i dividing G but not H. This completes the proof.

We now consider equations

$$g = dx^2 - bd'y^2 = N, \tag{10}$$

where $b = a$ or $2a$ according as $e = 0$ or 1, with aw^{-1} prime to $2H$ and N is prime to H. In order that (10) be solvable, it is necessary that $c_{p_i}(g) = (-N, bH)_{p_i}$ for all primes p_i, where $c_p(g)$ is the Hasse symbol (see [3], p. 47). We now choose $b \pmod{H}$, so that, for any fixed N prime to H, there is one and only one value of d such that $c_{p_i}(g) = (-N, bH)_{p_i}$ for all primes p_i dividing H.

THEOREM 6. *Let $p_1, p_2, \cdots, p_n$ be the prime factors of H. We can choose $b \pmod{p_i}$ $i = 1, 2, \cdots, n$ so that, for any N prime to H, there will be one and only one d dividing H such that*

$$c_{p_i}(g) = (-N, bH)_{p_i} \tag{11}$$

for all primes p_i dividing H.

Proof. First compute $c_p(g)$ and find it equal to $(-1, bH)_p (d, bH)_p = (-d, bH)_p$. Suppose $(-d_1, bH)_p = (-d_2, bH)_p$ for all such p. This implies and is implied by $(d_1 d_2, bH)_p = 1$ for such p. Hence our theorem will be proved if we can show that, for a proper choice of b, $(d, bH)_p = 1$ for all p dividing H holds for one and only one value of d (namely $d = 1$).

In order to prove this by induction, assume first that H is a prime; then the two possible values of d are 1 and p. Since $(1, bp)_p = 1$, we want to choose b so that $(p, bp)_p = -1$. Now $(p, bp)_p = (p, -b)_p = (-b \mid p)$. Hence choose b so that the latter is -1.

Assume the theorem for H containing prime factors $p_1, p_2, \cdots, p_{n-1}$ and that b has been chosen so that $(d, bH)_p = 1$ for all primes dividing H only when $d = 1$. In order to establish the theorem for $p_n H$ choose b' so that $(b' p_n \mid p_i) = (b \mid p_i)$ for $i = 1, 2, \cdots, n-1$. We must show that for $d \neq 1$ neither of the following is 1 for $p_i = p_1, p_2, \cdots, p_n$: $(d, b'Hp_n)_{p_i}$, $(dp_n, b'Hp_n)_{p_i}$.

First, $(d, b'Hp_n)_{p_i} = (d, bH)_{p_i}$ for $i = 1, 2, \cdots, n-1$ and since the latter is -1 for some i, the former is also. Second, $(dp_n, b'Hp_n)_{p_i} = (dp_n, bH)_{p_i}$ for $i = 1, 2, \cdots, n-1$ and the latter symbol is equal to $(d, bH)_{p_i}(p_n, bH)_{p_i}$; thus, if $(dp_n, b'Hp_n)_{p_i} = 1$ for $i = 1, 2, \cdots, n$ we have $(d, bH)_{p_i} = (p_n, bH)_{p_i}$ for $i = 1, 2, \cdots, n-1$. Thus, by the hypothesis of our induction, there is one and just one d, i.e., d_0, which has this property. Then, choose b' $(\operatorname{mod} p_n)$ so that $(d_0p_n, b'Hp_n)_{p_n} = -1$. This is possible since $(d_0p_n, b'Hp_n)_{p_n} = (d_0, b'Hp_n)_{p_n}(p_n, b'Hp_n)_{p_n} = (d_0 | p_n)(-b'H | p_n)$. Since for $d = 1$, $c_p(g) = (-1, bH)_p$ for all primes dividing H, the above shows that we need not consider (10) with $N = 1$ for $d \neq 1$. Theorem 5 shows that proper choice of a $(\operatorname{mod} G)$ also excludes $N = 1$, $d = 1$. This completes the proof and we complete the exclusion of (10) by the following theorem.

THEOREM 7. *There exist odd numbers* α, β *incongruent* $(\operatorname{mod} 4)$, *a divisor* d_0 *of* H, *and a choice of sign* $\pm$ *so that* (10) *with* $N = -1, 2, -2$ *have no solutions for the following conditions imposed upon* a:

(1) *If* $e = 0$, $aH \equiv 1 \pmod 4$ *or* $aH \equiv 5 \pm 2 \pmod 8$,

(2) *If* $e = 1 = w$, $aH \equiv \alpha$ *or* $\beta \pmod 8$,

(3) $a = wpq$ *where* p *and* q *are distinct primes not dividing* $2H$ *and* $(-d_0 | p) = -1$.

Proof. If $e = 0$, we may assume from the previous theorem that a is chosen $(\operatorname{mod} K)$ so that for all p_i dividing H each of the following three equations is solvable in $R(p_i)$ for only one divisor d of H:

$$(12)\quad dx^2 - a(H/d)y^2 = -1,\quad dx^2 - a(H/d)y^2 = +2,\quad dx^2 - a(H/d)y^2 = -2;$$

that is,

$$(13)\quad (-d, aH)_{p_i} = 1,\ (2d, aH)_{p_i} = 1,\ (-2d, aH)_{p_i} = 1 \text{ for all } p_i \text{ dividing } H.$$

Let the respective values of d be d_1, d_2, d_3. First, if $aH \equiv 1 \pmod 4$, the last two equations of (12) have no solution and we may exclude the first by choosing p so that $(-d_1, aH)_p = (-d_1 | p) = -1$. Second, consider $aH \equiv 3 \pmod 4$. Now $(-d_1, aH)_{p_i} = (2d_2, aH)_{p_i} = 1$ implies $(-2d_1d_2, aH)_{p_i} = 1$ for all p_i dividing H. Since there is just one divisor d_3 of H such that $(-2d_3, aH)_{p_i} = 1$ for all p_i dividing H, d_1d_2 is, except for a square factor, equal to d_3. Similarly, any product of two of d_1, d_2, d_3 is, except for a square factor, equal to the third. This implies that at least one of d_1, d_2, d_3 is congruent to $1 \pmod 4$. Suppose $-d_1 \equiv -1 \pmod 4$, then $aH \equiv 3 \pmod 4$ implies $(-d_1, aH)_2 = -1$ and one of $(2d_2, aH)_2$, $(-2d_3, aH)_2$ is -1. Suppose that it is the former. Then choose p so $(-2d_3, aH)_p = (-2d_3 | p) = -1$.

If $e = 1 = w$, we have the same set of equations except that a is replaced by $2a$. There are two odd numbers α such that $(2d_2, 2\alpha)_2 = (2d_2, 2)_2 (2d_2, \alpha)_2 = -1$ and the two are incongruent (mod 4). Make $aH \equiv \alpha \pmod 8$ and see that two of the equalities in (13) are denied for $p_i = 2$ and the third may be denied for $p_i = p$.

If $e = 1$, $w = 2$ equation (10) becomes $N = dx^2 - 4pqd'y^2$ which is not solvable for $N = 2$ or -2. The other equation is excluded by $(-d_0 | p) = -1$ where d_0 is the value of d for $N = -1$.

Theorems 5, 6 and 7 list restrictions imposed on p and q (mod $8K$) which assure the satisfaction of condition (i) of Theorem 3. We now consider condition (ii) of Theorem 3.

THEOREM 8. *Let h and h' be binary forms of the same primitive genus and of determinant JM; let d_0 be some integer prime to $2JM$ and $(JM, K) \leqq 2$; let h represent an integer a_0 satisfying the conditions* (mod $16K$) *imposed by Theorems* 5, 6 *and conditions* 1 *and* 2 *of Theorem* 7. *Then there are two primes p and q such that $(-d_0 | p) = -1$, wpq is represented by h and h', and $wpq \equiv a_0$* (mod $16K$).

Proof. We carry through the proof for h properly primitive. If h and h' are improperly primitive the same proof holds if we replace h and h' by $\frac{1}{2}h$ and $\frac{1}{2}h'$. This of course replaces JM by $\frac{1}{4}JM$.

The form hh' derived by composition from h and h' is in the principal genus and hence arises by duplication (cf. [3], p. 167); hence $hh' = v^2$ and, denoting the class vh'^{-1} by u, we have $h' = vu^{-1}$. Since d_0 is an odd factor of K and JM is the determinant of v with $(K, JM) \leqq 2$, there is an integer p' prime to $2 | f | M$ such that $(-d_0 | p') = -1$ and p' is compatible with the characters of v. For any a compatible with the characters of h and prime to JM we can find an integer q' prime to $2 | f | M$ so that $p'q' \equiv a \pmod m$ for any preassigned m prime to p'. Then v represents infinitely many primes $p \equiv p' \pmod{8 | f | M}$ and thus u and u^{-1} infinitely many primes $q \equiv q' \pmod{8 | f | M}$. Let p, q be one such pair and have $h = pq$, $h' = pq$ both solvable.

The preceding theorems combine to establish the following result.

THEOREM 9. *If an indefinite primitive ternary form f represents primitively a primitive form $-h$ of determinant JM with $(JM, K) \leqq 2$, then it represents every binary form in the same genus as $-h$ provided*

(i) *M is an odd prime or twice a prime according as the reciprocal form of f is properly or improperly primitive,*

(ii) *h represents an integer a satisfying the conditions* (mod $8K$) *of Theorems* 5, 6 *and* 7.

Then in view of this theorem and Theorem 2, the proof of Meyer's theorem will be completed if we establish

THEOREM 10. *If f and f' are two indefinite primitive ternary forms of the same genus, if $(J, K) \leqq 2$ and neither J nor K is divisible by 4, then they represent, respectively, two primitive binary forms $-h$ and $-h'$ having the following properties:*

(i) *h and h' are the same genus and have determinant JM,*

(ii) *M is a prime not dividing $2JK$ or twice such a prime according as the reciprocal form of f is properly or improperly primitive,*

(iii) *h represents an integer a satisfying conditions* 1 *and* 2 *of Theorem* 7.

Theorems 5 and 6 need not be considered here since $(JM, K) \leqq 2$, implies that h represents integers $\equiv a \pmod p$ for any odd prime p dividing K.

The proof of this theorem requires two lemmas.

LEMMA 1. *Let the invariants, J, K and w, of a primitive indefinite ternary form f be such that $(J, K) \leqq 2$, $J \not\equiv 0 \not\equiv K$* (mod 4) *and $w = 1$ or 2 according as the form f is properly or improperly primitive and, except for the factor 2, M is prime to JK. Then f represents a binary form $-h$ such that $h = a$ is solvable for a satisfying conditions* 1 *and* 2 *of Theorem* 7 *if M satisfies the following conditions exhibited in tabular form, where p. p. means properly primitive and i. p. means improperly primitive.*

Invariants	Conditions on M	Numbers represented (mod 8) by h
1. f p. p., J odd, K even	$M \equiv -J \pmod 4$ and $c_2(f) = (2 \mid -JM)$	arbitrary
2. F p. p., K odd, J even	arbitrary $JM \equiv \mp 2c_2(f) \pmod 8$	$a \equiv H \pmod 4$ and $c_2(f) = (2 \mid aH)$ $aH \equiv 5 \pm 2 \pmod 8$
3. f i. p., J odd, K even	arbitrary	$a \equiv 2$ or $6 \pmod 8$
4. F i. p., K odd, J even	$M \equiv J \pmod 8$	$a \equiv K \pmod 4$
5. JK odd	$c_2(f) = 1$ and $M \equiv 3J \pmod 8$ $c_2(f) = -1$ and $M \equiv J \pmod 4$	arbitrary $a \equiv -H \pmod 4$
6. $J \equiv K \equiv 2 \pmod 4$	$c_2(f) = (2\alpha, -JM)_2$ $c_2(f) = (2\beta, -JM)_2$	$a \equiv \alpha H \pmod 8$ $a \equiv \beta H \pmod 8$

(α and β defined in Theorem 7)

Proof. Since there are no requirements imposed by Theorem 7 for f improperly primitive, we need consider only properly primitive forms f in this proof. First we show that under the conditions imposed, f represents a primitive binary form $-h$ of determinant JM. Let g be the greatest odd factor common to all the coefficients of h. Then $g^2|J$ and, since J and M have no odd factors in common, we may by an equivalence transformation on F have

$$F \equiv cJKx^2 + bKy^2 + Mz^2 \pmod{J_0^2},$$

where J_0 is the greatest odd factor of J and bc is prime to J_0. Then

$$f \equiv bMx^2 + cJMy^2 + bcJKz^2 \pmod{J_0^2},$$

and the leading binary of f has determinant JM and leading coefficient congruent $(\operatorname{mod} g^2)$ to bM which is therefore prime to g. Thus the coefficients of this leading binary form have no odd prime factor in common. This shows that h may be taken to be primitive.

To show that h is properly primitive notice first that it can be improperly primitive only if $JM \equiv 3 \pmod 4$ or $4 \pmod 8$. We postpone the latter and consider the former. If, under this condition, it is improperly primitive, it can be considered to have a leading element $-2s \equiv 2 \pmod 4$. Then

$$c_2(f) = (-2sK, -JM)_2. \tag{14}$$

If K is odd, then by the conditions imposed on M either $c_2(f) = 1$ in which case $JM \equiv 3 \pmod 8$ implies $(-2sH, -JM)_2 = -1 \neq c_2(f)$, or $c_2(f) = -1$ and $JM \equiv 1 \pmod 4$. Thus, for K odd, condition 5 of the lemma implies that f represents no improperly primitive forms of determinant JM. If $K \equiv 2 \pmod 4$ and f represents an improperly primitive binary h of determinant JM, then f may be taken congruent to $h + cx_3^2 \pmod{2^r}$ with c even, which denies f properly primitive. It remains to consider the numbers which h represents and $JM \equiv 4 \pmod 8$.

If $JM \not\equiv 0 \pmod 4$, the numbers $(\operatorname{mod} 8)$ which h represents will be determined by its Hasse symbol with $p = 2$. Furthermore h represents a $(\operatorname{mod} 8)$ if and only if $c_2(f) = (aK, -JM)_2$. Now consider the conditions in order.

1. Here $M \equiv -J \pmod 4$ shows that $c_2(f) = (2 | -JM)$ and all odds $(\operatorname{mod} 8)$ are represented by h.

2. Here the numbers a represented by h are given by $c_2(f) = (aH, -JM)_2$. Thus if $aH \equiv 1 \pmod 4$, a is represented if and only if $c_2(f) = (aH, -JM)_2$

$= (2|aH)$ regardless of the choice of M. If $aH \equiv 5 \pm 2 \pmod 8$ the choice of $M \pmod 4$ specified suffices.

3. If f is improperly primitive, Theorem 7 imposes no restrictions.

4. If $M \equiv J \pmod 8$, h represents numbers congruent to $K \pmod 4$ and is properly primitive, for F may be taken congruent to

$$g + 2cx_3^2 \pmod{2^r},$$

where g is an improperly primitive binary form. Kf is then congruent to $2g_1 + ex_3^2 \pmod{2^r}$, where $e \equiv -1 \pmod 4$, and g_1 is an improperly primitive binary form. However g_1 cannot have determinant $JM/4$ since the condition imposed requires $JM/4 \equiv 1 \pmod 4$. This shows that f does not represent an improperly primitive binary form of determinant $JM/4$. On the other hand, the odd numbers represented by Kf are congruent to e, that is, congruent to $-1 \pmod 4$. This shows that the odd numbers represented by h are congruent to $K \pmod 4$.

5. If $c_2(f) = 1$ and $JM \equiv 3 \pmod 8$, then $c_2(f) = (aH, -JM)_2$ for a arbitrary. But $c_2(f) = -1$ and $JM \equiv 1 \pmod 4$ shows that $-1 = (aH, -1)_2$ and hence $aH \equiv -1 \pmod 4$.

6. Here $a \equiv \alpha H \pmod 8$ if and only if $c_2(f) = (2\alpha, -JM)_2$. This completes the proof.

Our final step in the proof of Theorem 10 and hence of Meyer's theorem is

LEMMA 2. *If the invariants J, K and w satisfy the conditions of the theorem and if f and f' are indefinite forms of the same primitive genus with these invariants, then F and F', their reciprocal forms, represent a prime M or twice such a prime satisfying the conditions imposed on M in Lemma* 1.

Proof. Since the classes f and $-f$ are class isomorphic, for this lemma we take J and K to be negative integers. This makes the binary h and the binary g, determined below, indefinite forms and therefore capable of representing both positive and negative integers.

If the form f represents wq where q is a prime not dividing $2|f|$ then, by Theorem 9 with f and F interchanged, F represents a binary form $-g$ of determinant qK and all binaries in the genus of $-g$ and hence all numbers $-r$, where r is consistent with the generic characters of g, provided conditions 1 and 2 of Theorem 7 hold after replacing a, J and K by $-r$, K and J respectively, that is

1′) If J is odd, $-rJ \equiv 1 \pmod 4$ or $-rJ$ congruent to a preassigned one of 3, 7 (mod 8).

2′) If J is even, F properly primitive, $-rJ/2 \equiv \alpha_1$ or $\beta_1 \pmod 8$ for α_1 and β_1 preassigned with $\alpha_1\beta_1 \equiv -1 \pmod 4$.

By Lemma 1 with f and F, a and $-r$ interchanged, the representation of such an r by F will be assured by the following choices of wq exhibited in tabular form. (Note that $c_2(f) = c_2(F)$.)

Invariants	Conditions on q	Conditions (mod 8) on r
f p. p., J odd, K even	arbitrary	$-r \equiv J \pmod 4$ and $c_2(f) = (2 \mid -rJ)$
	$Kq \equiv \mp 2c_2(f) \pmod 8$	$-rJ \equiv 5 \pm 2 \pmod 8$
F p. p., J even, K odd	$q \equiv -H \pmod 4$ and $c_2(f) = (2 \mid -qH)$	arbitrary
f i. p., J odd, K even	$q \equiv H \pmod 4$	$-r \equiv J \pmod 4$
F i. p., J even, K odd	arbitrary	$r \equiv 2$ or $6 \pmod 8$
JK odd	$c_2(f) = 1$ and $q \equiv 3H \pmod 8$	arbitrary
	$c_2(f) = -1$ and $q \equiv H \pmod 4$	$-r \equiv -J \pmod 4$
$J \equiv K \equiv 2 \pmod 4$	$c_2(f) = (2\alpha_1, -Kq)_2$	$-r \equiv \alpha_1 J/2 \pmod 8$
	$c_2(f) = (2\beta_1, -Kq)_2$	$-r \equiv \beta_1 J/2 \pmod 8$

First, to show that such a choice of wq is possible, notice ([3], p. 87), f properly primitive implies that f represents all odds (mod 8) unless K is odd and $c_2(f) = -1$. In this case the only odds excluded (mod 8) are those numbers b for which $bH \equiv -1 \pmod 8$. Inspection of the table shows that choice of q may be made consistent with this exclusion. Similarly if f is improperly primitive $2q$ is excluded (mod 8) only when $c_2(f) = -1$ and $qH \equiv -1 \pmod 8$, where the corresponding entry in the table is $q \equiv H \pmod 4$.

Next, comparison of the listing for r in the above table with the requirements on M imposed by Lemma 1 shows that they are consistent except for the last case which needs further consideration. The only possibility of trouble for this case is when none of

$$c_2(f) = (2\alpha, -JM)_2, \qquad c_2(f) = (2\beta, -JM)_2$$

hold for M $-\alpha_1 J/2$ or $-\beta_1 J/2 \pmod 8$; that is, all of

$$(2\alpha, +2\alpha_1)_2, \quad (2\alpha, +2\beta_1)_2, \quad (2\beta, +2\alpha_1)_2, \quad (2\beta, +2\beta_1)_2$$

are equal to $-c_2(f)$. This is impossible since their product is equal to $(\alpha\beta, \alpha_1\beta_1)_2$ which has the value -1.

Thus we have shown that F represents a prime M_1 or double such a prime consistent with the requirements of Lemma 1. In similar fashion f' may be shown to represent a prime q' congruent to q for an arbitrary modulus. Then F' represents a binary form $-g'$ which represents a prime $-M_2$ congruent to M_1 (mod $8K$) since the determinant of g' has at most a factor of 2 in common with K. Then a prime $M \equiv M_1 \pmod{q}$ and $\equiv M_2 \pmod{q'}$ as well as congruent to M_1 (mod $8K$) will be represented by both F and F'. This completes the proof.

THE UNIVERSITY OF COLORADO.

BIBLIOGRAPHY.

[1] L. E. Dickson, *Studies in the theory of numbers*, University of Chicago Press, 1930.

[2] M. Eichler, *Quadratische Formen und Orthogonale Gruppen*, Berlin, 1952.

[3] B. W. Jones, *The Arithmetic Theory of Quadratic Forms*, New York, 1950.

[4] A. Meyer, *Zur Theorie der unbestimmten ternären quadratischen Formen*, Zürich, 1871.

[5] ———, "Zur Theorie der Indefiniten Ternären Quadratischen Formen," *Journal für Mathematik*, vol. 108 (1891), pp. 125-139.

ON THE BEHAVIOR OF INVARIANT CURVES NEAR A HYPERBOLIC POINT OF A SURFACE TRANSFORMATION.*

By SHLOMO STERNBERG.

1. In [2], Poincaré considered surface transformations of the type

$$x_1 = sx + f(x, y), \qquad y_1 = ty + g(x, y), \tag{1}$$

where s and t are real with $s > 1 > t > 0$, and where f and g are real analytic functions of x and y which vanish at the origin together with their first derivatives. He proved, by the majorant method, that there exist two analytic invariant curves each tangent to one of the axes. In the case of dynamics, the relation $st = 1$ holds (as a necessary condition for the transformation to be area preserving). We shall not make use of this condition. Hadamard [1], treating the case where f and g are merely assumed to be of the class C^1 (and to vanish at the origin along with their first derivatives), showed the existence of invariant curves by the method of successive approximations. In paragraphs 2-6 we shall treat questions of existence and uniqueness of invariant curves by a geometric method under slightly less restrictive conditions and show that our results are, in a sense, the best possible. In paragraphs 7 and 8 we shall deal with questions of smoothness of invariant curves and show that a C^n assumption on the non-linear terms implies that the invariant curves are of class C^n $(n \geqq 1)$. The problems treated in this paper were suggested to the author by Professor Wintner.

2. Let T be a transformation of type (1) which is topological in some neighborhood of the origin. Let $s > 1 > t > 0$, and let the functions f, g be continuous and $o(r)$ as $r = (x^2 + y^2)^{\frac{1}{2}} \to 0$. It is then clear that for all points sufficiently close to the origin, the y coordinate is decreased and the x coordinate is increased upon application of T. Thus if an arc passing through the origin is to be invariant, it must be tangent to one of the axes or else experience oscillations of increasing amplitude and frequency near the origin. In the latter case we shall, naturally, not speak of an "invariant curve." Hence, either the arc is tangent to the x-axis at the origin, in which case all points of the arc move in toward the origin upon application of T; or

---

* Received February 7, 1955.

the arc is tangent to the y-axis, in which case all the points move away from the origin.

3. In this paragraph and the following, S and T will denote transformations of type (1) where $t > 1$. $t > s$ and $s > 1$, $s > t$ respectively.

THEOREM 1. *Let S be a transformation of type* (1) *where* $0 < s < 1$. $s < t$ *and where f, g are continuous and are $o(r)$ as* $r = (x^2 + y^2)^{\frac{1}{2}} \rightarrow 0$. *Then there exist a neighborhood N of the origin and a closed set E in N such that* 1) *E is invariant under S,* 2) *for every abscissa x in N there exists a point (x, y) of E,* 3) *E is tangent to the x-axis of the origin* (i.e., given any $\epsilon > 0$, for all sufficiently small $|x|$, all points (x, y) of E lie in the angular region $|y| \leqq \epsilon |x|$).

Proof. We can restrict our attention to the right half plane $x \geqq 0$ since the argument is the same for the left half plane. Consider a point (x, y) on the line $y = \alpha x$, $\alpha > 0$. Its image will be the point (x_1, y_1), where $x_1 = sx + f(x, y)$, $y_1 = ty + g(x, y)$. Thus

$$y_1/x_1 = \{\alpha t/s + g(x, y)/sx\}/\{1 + f(x, y)/sx\}.$$

Since $y = \alpha x$ and both f and g are $o(r)$, we find that $x_1 < x$ and $y_1 > \alpha x_1$ for all sufficiently small positive x, say, for $0 < x \leqq x_0$.

In other words, a point on the line $y = \alpha x$ is moved to the left and above the line by S for all $x < x_0$. Similarly, a point on the line $y = -\alpha x$ is moved to the left and below the line. For all $x < x_0$, let $E^0(x)$ denote the vertical line segment joining the points $(x, -\alpha x)$ and $(x, \alpha x)$. Now S will map $E^0(x)$ into a set containing a connected arc joining a point above the wedge $|y| \leqq \alpha x$ to a point below the wedge. Let $E^1(x)$ denote the subset of $E^0(x)$ consisting of those points whose images under S lie in the wedge $|y| \leqq \alpha x$. If $E^0(x), E^1(x), \cdots, E^{n-1}(x)$ have been defined, let $E^n(x)$ denote the subset of $E^{n-1}(x)$ consisting of those points whose images under S^n lie in the wedge. Since, by induction, the image of $E^{n-1}(x)$ under S^n contains a connected arc joining a point above the wedge to a point below it, $E^n(x)$ is not empty. $E^n(x)$ is clearly closed for all n. Hence $E(x) = \bigcap_n E^n(x)$ is a non-empty closed subset of $E^0(x)$. It is clear that given any ϵ we can find a positive number ξ such that for all $0 < x \leqq \xi$ and all $k \geqq \epsilon$, S takes a point on the line $y = kx$ into a point above the line $y = (k + \tau)x$ where τ is independent of x and k. Hence for a point (x, y) with $x \leqq \xi$ to lie in E we must have $|y| < \epsilon x$. Thus $E(x)$ is tangent to the x-axis at the origin.

Let $E = \bigcup_x E(x)$. Then E is clearly invariant under S. In fact, any subset of the wedge invariant under S is eventually contained in E. Furthermore, a point $P = (x, y)$ of the closed wedge $|y| \leqq \alpha x$, $0 \leqq x \leqq x_0$ is not in E only if S^n, for some positive n, carries P outside the closed wedge. But then a small neighborhood about P is carried outside the wedge by S^n. Thus the complement of E is open, so that E is closed. This proves 1), 2) and 3).

THEOREM 2. *Let T be a transformation of type* (1) *where $s > 1$, $s > t$ and f and g are continuous and $o(r)$ as $r \to 0$. Then there exist a neighborhood N of the origin and a closed set E satisfying* 1), 2), *and* 3) *of Theorem* 1 *where, by invariance, we mean $T(E) \cap N = E$.*

Remark. If T is topological, then Theorem 2 can be obtained by applying Theorem 1 to $S = T^{-1}$.

In order to prove Theorem 2, let us consider an angular wedge W: $|y| \leqq kx$. Then from arguments similar to the above we have

$$(T(W) \cap N) \subset (W \cap N).$$

Let

$$F_0 = W \cap N, \qquad F_1 = T(W \cap N) \cap N$$

and in general $F_n = T(F_{n-1}) \cap N$. It is clear, by induction, that $F_{n+1} \subset F_n$ and thus $E = \bigcap_n F_n$ is a non-empty closed set. It is clear from arguments similar to those of the preceding theorem that 1), 2) and 3) are satisfied.

4. We shall now show that, under the hypotheses of Theorem 1, $E(x)$ can contain more than one point. S will be a transformation of type (1) in which $f \equiv 0$. We shall choose g such that $g(x, 0) = 0$ and so that the x-axis is an invariant curve.

Let $y = \psi(x)$ be a smooth curve tangent to the x-axis at the origin and such that $\psi(x) > 0$ for $x > 0$. In order that $\psi(x)$ be an invariant curve, it is necessary and sufficient that the functional equation

$$t\psi(x) + g(x, \psi(x)) = \psi(sx) \tag{2}$$

be satisfied. Rewriting this we obtain the condition on g

$$g(x, \psi(x)) = \psi(sx) - t\psi(x). \tag{2'}$$

We now define g as follows: Let $g(x, y)$ be 0, $\{\psi(sx) - t\psi(x)\}y/\psi(x)$ or $\psi(sx) - t\psi(x)$ according as $y \leqq 0$, $0 \leqq y \leqq \psi(x)$ or $y \geqq \psi(x)$. We must now verify that the mapping S, with $f \equiv 0$ and g defined as above, satisfies the conditions of Theorem 1. (It is clear that ψ can be so chosen that the mapping is even topological.) Now $|g(x,y)| \leqq \psi(sx) + t\psi(x)$. Thus g is $o(|x|)$ and so *a fortiori* $o(r)$.

We have constructed g so that (2′) is satisfied; hence, both $y = \psi(x)$ and $y = 0$ are invariant curves of S.

5. We now consider the question of uniqueness.

THEOREM 3. *If the hypotheses of Theorem 1 are satisfied, and if both f and g satisfy uniform Lipschitz conditions with respect to x and y, where in a neighborhood of the origin, the uniform Lipschitz constants are less than $\delta < \frac{1}{4}(t - s)$, then, in a suitably small neighborhood N, E is a curve $y = \psi(x)$. In other words, there exists a unique invariant arc in any wedge $|y| \leqq k|x|$ for sufficiently small $|x|$.*

In order to prove this theorem it will suffice to show that given any wedge $|y| \leqq kx$ there exists a positive number ξ such that there cannot be two points (x,y) and (x,y'), where $0 < x \leqq \xi$ and $y > y'$, whose images under S^n are in $|y| \leqq kx$ for all non-negative n.

Since f is $o(r)$, any point in $|y| \leqq kx$ with sufficiently small abscissa $x\,(> 0)$ is mapped into a point whose abscissa is smaller than $(s + \epsilon)^n x$. Thus if (x_n, y_n) and (x'_n, y'_n) denote the images under S^n of (x,y), (x,y'), we must have x_n, $x'_n < (s + \epsilon)^n x$ and consequently, $|y_n - y'_n| < 2k(s + \epsilon)^n x$ in order that these images all lie in $|y| \leqq kx$. We shall show that this is impossible for all suitably small x.

From (1), $x_n - x'_n = s(x_{n-1} - x'_{n-1}) + f(x_{n-1}, y_{n-1}) - f(x'_{n-1}, y'_{n-1})$. Thus, for all sufficiently small x,

$$|x_n - x'_n| \leqq s\,|x_{n-1} - x'_{n-1}| + \delta\,|x_{n-1} - x'_{n-1}| + \delta\,|y_{n-1} - y'_{n-1}|.$$

Similarly, $y_n - y'_n = t(y_{n-1} - y'_{n-1}) + g(x_n, y_n) - g(x'_n, y'_n)$; so that

$$|y_n - y'_n| > t\,|y_{n-1} - y_{n-1}| - \delta\,|x_{n-1} - x'_{n-1}| - \delta\,|y_{n-1} - y'_{n-1}|.$$

Now assume that

$$(3) \qquad |x_{n-1} - x'_{n-1}| < |y_{n-1} - y'_{n-1}|;$$

then we have

$$|x_n - x'_n| < (s + 2\delta)|y_{n-1} - y'_{n-1}| \quad \text{and} \quad |y_n - y'_n| > (t - 2\delta)|y_{n-1} - y'_{n-1}|.$$

Since $\delta < \frac{1}{4}(t - s)$ we obtain $|x_n - x'_n| < |y_n - y'_n|$. As $|x_0 - x'_0| = 0 < |y_0 - y'_0|$, the inequality (3) is true for all n, by induction. We thus obtain

$$|y_n - y'_n| > (t - 2\delta)|y_{n-1} - y'_{n-1}| \quad \text{or} \quad |y_n - y'_n| > (t - 2\delta)^n\,|y - y'|.$$

We must therefore have $(t-2\delta)^n|y-y'|<2k(s+\epsilon)^n x$, which is clearly impossible for all n if ϵ is chosen to be less than $\frac{1}{2}(t-s)$. Thus $y=y'$ and the theorem is established.

6. Let us now consider a transformation T satisfying the hypotheses of Theorem 2. The set E, in virtue of its construction, contains any invariant set lying in the intersection of any angular wedge $|y|\leqq kx$ and some sufficiently small neighborhood of the origin. Now let G_0 be any set lying entirely in such a region. It follows, from reasoning similar to that of paragraph 3, that the image of this set under T, G_1, still lies in the angular region $|y|\leqq kx$. Let $G_0, G_1, G_2, \cdots, G_n$ be the sequence of sets so obtained. Let C be the set of limit points of this sequence. C is obviously an invariant set and hence a subset of E. In particular, if E reduces to a curve, then all the successive approximations converge to the invariant curve. Let us now assume that both f and g satisfy uniform Lipschitz conditions (with respect to x and y) where the Lipschitz constant can be made arbitrarily small by restricting ourselves to a small neighborhood of the origin. Let $y=\phi_0(x)$ be a curve satisfying a uniform Lipschitz condition with Lipschitz constant K. Thus $|(y-y')/(x-x')|<K$. Using the same notation as before, we obtain

$$(y_1-y'_1)/(x_1-x'_1)$$
$$=\{t(y-y')+g(x,y)-g(x',y')\}/\{s(x-x')+f(x,y)-f(x',y')\},$$

so that $|(y_1-y'_1)/(x_1-x'_1)|<\{tK+(1+K)\delta\}/\{s-(1+K)\delta\}$, where δ can be made arbitrarily small by choosing x and x' sufficiently small. Thus in a sufficiently small region about the origin, all the image curves satisfy a Lipschitz condition with Lipschitz constants all smaller than K. Hence the sequence is equi-continuous and the convergence is uniform. This is essentially the case considered by Hadamard [1]. We have thus proved

THEOREM 4. *Let T be a transformation of type* (1) *where $s>1$, $s>t$, f and g are continuous and $o(r)$. Furthermore let the set E of Theorem 2 be a curve. Then the sequence of successive images of a set situated in an angular region $|y|<kx$ converge for sufficiently small x. If, in addition, f and g satisfy uniform Lipschitz conditions with Lipschitz constants which are $o(1)$ as $r\rightarrow 0$, then the iterates of any curve $y=\phi(x)$ which satisfies a uniform Lipschitz condition converge uniformly for sufficiently small x to the invariant curve, which then must satisfy a uniform Lipschitz condition.*

7. We shall now consider differentiability of the invariant curve.

THEOREM 5. *Let S be a transformation of type (1), where $0 > s > t$, $s > 1$ and where f and g are of class C^1 and vanish together with their first derivatives at the origin. Then the unique invariant curve tangent to the x-axis at the origin is of class C^1 for all sufficiently small x.*

In order to make the method clear, we shall first go through the proof for the particular case when $f(x, y) \equiv 0$. Let $y = \psi(x)$ be the invariant curve. It must then satisfy the functional equation (2). Writing this equation for two distinct points x and x' and subtracting, we obtain

$$t\psi(x) - t\psi(x') + g(x, \psi(x)) - g(x', \psi(x')) = \psi(sx) - \psi(sx').$$

Applying the mean value theorem, we obtain

$$t\{\psi(x) - \psi(x')\}/(x - x') + g_x(\xi^*, \psi(x)) + g_y(x', \psi(\xi))\{\psi(x) - \psi(x')\}/(x - x')$$
$$= \{\psi(sx) - \psi(sx')\}/(x - x') \text{ or}$$

$$\{\psi(x) - \psi(x')\}/(x - x')$$
$$= [t + g_y(x', \psi(\xi))]^{-1}[-g_x(\xi^*, \psi(x)) + s\{\psi(sx) - \psi(sx')\}/(sx - sx')],$$

where both ξ and ξ^* lie between x and x'.

We have thus expressed $\{\psi(x) - \psi(x')\}/(x - x')$ in terms of

$$\{\psi(sx) - \psi(sx')\}/(sx - sx').$$

Repeating the process n times, we obtain

$$\{\psi(x) - \psi(x')\}/(x - x')$$
$$= -\sum_{k=1}^{n} s^{k-1} g_x(\xi^*_{k-1}, \psi(s^{k-1}x))\{\prod_{j=1}^{k}(t + g_y(s^{j-1}x, \psi(\xi_{j-1}))\}^{-1}$$
$$+ s^n\{\prod_{j=1}^{n}(t + g_y(s^{j-1}x, \psi(\xi_{j-1}))\}^{-1}\{\psi(s^n x) - \psi(s^n x')\}/(s^n x - s^n x'),$$

where both ξ_j and ξ^*_j lie between $s^j x$ and $s^j x'$.

Now $\psi(x)$ satisfies a uniform Lipschitz condition (by Theorem 4) and $|g_y| < t - s - \delta$ for all sufficiently small x and x'. Thus the second term in the above equation tends to zero and we obtain

$$\{\psi(x) - \psi(x')\}/(x - x') = \sum_{k=1}^{\infty} s^k g_x(\xi^*_{k-1}, \psi(s^{k-1}x)) \prod_{j=1}^{k}(t + g_y(s^{j-1}x, \psi(\xi_{j-1})))^{-1}.$$

Since g_x and g_y are $o(1)$ as $x \to 0$, this series is majorized by a series of the form $\sum_k s^{k-1}(t-\epsilon)^{-k}$ and so is uniformly convergent. We can, therefore, let $x \to x'$ and obtain

$$\psi'(x) = \sum_{k=1}^{\infty} s^{k-1} g_x(s^{n-1}x, \psi(s^{k-1}x)) \prod_{j=1}^{k} (t + g_y(s^{j-1}x, \psi(s^{j-1}x)))^{-1}.$$

In the case where f is not identically zero the iteration is more difficult. The functional equation satisfied by the invariant curve $\psi(x)$ now takes the form

$$t\psi(x) + g(x, \psi(x)) = \psi(sx + f(x, \psi(x))).$$

(This is the functional equation used by Poincaré [2]). Proceeding as before,

$$t\{\psi(x) - \psi(x')\} + g(x, \psi(x)) - g(x', \psi(x')) = \psi(x_1) - \psi(x'_1),$$

where $(x_1, \psi(x_1))$ and $(x'_1, \psi(x'_1))$ are the images, respectively, of $(x, \psi(x))$ and $(x', \psi(x'))$. Thus

$$\begin{aligned} t\{\psi(x) - \psi(x')\}/(x - x') + g_x(\xi^*, \psi(x)) + g_y(x', \psi(\xi))\{\psi(x) - \psi(x')\}/(x - x') \\ = [\{\psi(x_1) - \psi(x'_1)\}/(x_1 - x'_1)][s + f_x(\eta^*, \psi(x)) \\ + f_y(x', \psi(\eta))\{\psi(x) - \psi(x')\}/(x - x')], \end{aligned}$$

where ξ, ξ^*, η and η^* are all between x and x'. Hence, the iteration takes the form $[\ ]_0 = \{(s + \sigma_0)[\ ]_1 + \delta_0\}\{\epsilon_0[\ ]_1 + (t + \tau_0)\}^{-1}$, where

$$[\ ]_0 = \{\psi(x) - \psi(x')\}/(x - x'), \qquad [\ ]_1 = \{\psi(x_1) - \psi(x'_1)\}/(x_1 - x'_{11}),$$

$$\sigma_0 = f_x(\eta^*, \psi(x)), \quad \delta_0 = -g_x(\xi^*, \psi(x)), \quad \epsilon_0 = f_y(x', \psi(\eta)), \quad \tau_0 = g_y(x', \psi(\xi)).$$

Introducing corresponding notations for the k-th iterate of this relation, we obtain

$$[\ ]_0 = \prod_{j=0}^{k} \begin{pmatrix} s + \sigma_j & \delta_j \\ \epsilon_j & t + \tau_j \end{pmatrix} [\ ]_{k+1}, \text{ where } \begin{pmatrix} a & b \\ c & d \end{pmatrix}(z) = (az + b)/(cz + d),$$

and the product denotes matrix multiplication.

We now wish to establish the convergence of this iteration. We first show that for all sufficiently small x and x', the product

$$\prod_{j=0}^{k} \begin{pmatrix} s + \sigma_j & \delta_j \\ \epsilon_j & t + \tau_j \end{pmatrix}$$

can be put in the form

$$\begin{pmatrix} A_k & B_k \\ C_k & 1 + D_k \end{pmatrix}, \text{ where } A_k = A \prod_1^k \alpha_j,\ C_k = C \prod_1^k \gamma_j,$$

with $|\alpha_j| < r < 1$, $|\gamma_j| < r < 1$, and $B_k = B_{k-1} + \beta_k$, $D_k = D_{k-1} + \partial_k$, with

$|\beta_k| < p^k$, $|\partial_k| < p^k$, and $|p| < 1$. This is certainly true for the case $k = 0$. We shall prove that the matrix has this form for all n, by induction. Thus assume the proposition true for the case $k = n$. Then the $(n+1)$-st matrix is

$$\begin{pmatrix} A_n & B_n \\ C_n & 1 + D_n \end{pmatrix} \begin{pmatrix} s + \sigma_n & \delta_n \\ \epsilon_n & t + \tau_n \end{pmatrix}.$$

Now since this matrix product is considered as acting as a linear fractional transformation we may divide all four elements by the same amount without changing the effect of the matrix. Carrying out the matrix multiplication and dividing by $t + \tau_n$ we obtain the matrix

$$\begin{pmatrix} A_n(s+\sigma_n)\rho_n + B_n\epsilon_n\rho_n & A_n\delta_n\rho_n + B_n \\ C_n(s+\sigma_n)\rho_n + (1+D_n)\epsilon_n\rho_n & 1 + D_n + C_n\delta_n\rho_n \end{pmatrix},$$

where $\rho_n = 1/(t + \tau_n)$. Now by virtue of the inductive hypothesis, $|B_n| < B$ and $|D_n| < D$, where B and D do not depend on n. Since σ_n, δ_n, ϵ_n, τ_n are all uniformly $o(1)$ (as x and x' go to zero), we may choose x and x' so small that

$$|(s + \sigma_n + \epsilon_n B/A)/(t + \tau_n)| < r < 1,$$

$$|(s + \sigma_n + \epsilon_n(1 + D_n)/C)/(t + \tau_n)| < r < 1,$$

$$|A\delta_n/(t + \tau_n)| < p < 1, \qquad |C\delta_n/(t + \tau_n)| < p < 1.$$

Thus the $(n+1)$-st matrix has the desired form. It should be remarked at this point that the constants C and D are still at our disposal.

As before, we know that the $[\ ]_n$ are uniformly bounded. We now choose C and D so small that the denominator of the n-th iterate never vanishes and the uniform convergence of the iteration process is established.

COROLLARY. *If, in addition to the hypotheses of Theorem* 5, *f and g are assumed to be of class C^n, then ψ is of class C^n in some neighborhood of the origin.*

Proof. The function $\psi'(x)$ can be represented as the quotient of two series both of which are majorized by $\sum Kr^n$ with $r < 1$. In taking successive difference quotients, one obtains series majorized by $K\sum M^k(n!/(n-k)!)r^n$, which is also uniformly convergent.

8. Let T be a transformation of type (1), where f and g are of class C^n $(n \geqq 1)$ and vanish at the origin together with their first derivatives. Furthermore, let $s > 1 > t > 0$. Then both T and T^{-1} satisfy the conditions of the corollary to Theorem 5 and hence, there exist exactly two invariant

curves of the transformation T, each tangent to one of the axes at the origin. We can perform a rotation of the coordinate system so that now neither invariant curve touches an axis. In these new (u, v) coordinates the two invariant curves may be represented as $v = \phi_1(u)$ and $v = \phi_2(u)$. We now write $X = v - \phi_1(u)$, $Y = v - \phi_2(u)$. Since the curves ϕ_1 and ϕ_2 are perpendicular at the origin, the Jacobian of this transformation does not vanish. The transformation T is therefore defined in these new variables. We thus have

THEOREM 6. *Let T be a transformation of type* (1) *with $s > 1 > t > 0$. Furthermore, let f and g be of class C^n $(n \geqq 1)$ and vanish at the origin together with their first derivatives. Then there exists a non-singular change of coordinates $X = X(x, y)$, $Y = Y(x, y)$ of class C^n so that in the new coordinates the axes are the invariant curves and T has the form*

$$X_1 = s(X + F(X, Y)), \qquad Y_1 = t(Y + G(X, Y)),$$

where F and G are both of class C^n, and where

$$F(0, y) = G(x, 0) = F_x(0, 0) = F_y(0, 0) = G_x(0, 0) = G_y(0, 0) = 0.$$

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] J. Hadamard, *Selecta*, 1901, pp. 163-166.
[2] H. Poincaré, Œuvres, vol. 1 (1886), pp. 202 204.

ON A PROBLEM OF LITTLEWOOD.*

By R. SALEM.

In his "Research Problems" Littlewood raises the question to know whether it is true that, given n distinct integers $m_1, m_2, \cdots, m_n$, there exists an absolute constant A satisfying

$$\int_0^{2\pi} | e^{m_1 ix} + \cdots + e^{m_n ix} | \, dx > A \log n.$$

The purpose of this paper is to prove the following weaker result.

THEOREM. *Let* m_n, $n = 1, 2, \cdots$, *be an increasing sequence of positive integers such that* $\log m_n = O(\log n)$. *Then, as* $n \to \infty$,

$$\limsup \int_0^{2\pi} | \cos m_1 x + \cdots + \cos m_n x | \, dx / (\log n)^{\frac{1}{2}} > 0.$$

The proof is extremely indirect and is based on the refinement of an argument used by the author (see Salem [2]) to give a new proof of the theorem of Menchoff on the convergence almost everywhere of series of orthogonal functions: namely, if $\{\phi_j\}_1^\infty$ is any orthonormal system, and if $\sum \gamma_j^2 \log^2 j < \infty$, then $\sum \gamma_j \phi_j$ converges almost everywhere.

LEMMA I. *Given* $\Omega(n)$ *increasing to* ∞ *and such that* $\Omega(n) = o(\log^2 n)$, *there exists an orthonormal system* $\{\phi_j\}_1^\infty$ *in* $(0, 1)$ *and a sequence of coefficients* $\{\alpha_j\}_1^\infty$ *such that* $\sum \alpha_j^2 \Omega(j) < \infty$ *and that the series* $\sum \alpha_j \phi_j$ *diverges almost everywhere.*

This lemma which is equivalent to the rough statement that "Menchoff's theorem can not be improved" is known. (See Kaczmarz and Steinhaus [1].)

LEMMA II. *Given* $\Omega(n)$ *increasing to* ∞ *and such that* $\Omega(n) = o(\log^2 n)$, *there exist*

1) *a fixed orthonormal system* $\{\phi_j\}_1^\infty$ *in* $(0, 1)$,
2) *for each* n, *a non-increasing sequence* $\{F_j\}_1^n$ *of characteristic functions of sets in* $(0, 1)$ *(depending on* n*)*,

---

* Received December 27, 1954.

3) *for each n, a function $Q \varepsilon L^2$ in $(0,1)$ (depending on n), with* $\int_0^1 Q^2 dx = 1$, *such that*

$$\limsup \sum_1^n \left(\int_0^1 QF_j\phi_j\,dx\right)^2/\Omega(n) = \infty.$$

Proof. Take for $\{\phi_j\}_1^\infty$ the orthonormal system whose existence is asserted by Lemma I. If it were true that

$$\sum_1^n \left(\int_0^1 QF_j\phi_j\,dx\right)^2 < C\Omega(n), \tag{1}$$

C being a constant, no matter how the sequence $\{F_j\}_1^n$ and the function Q were chosen for each n, it would follow, by a known argument (see Salem [2]) that given any n numbers $\gamma_1, \cdots, \gamma_n$ and writing $S_p(x) = \sum_1^p \gamma_j\phi_j$, one would have

$$\int_0^1 \sup_{1\leqq p\leqq n} |S_p(x)|^2\,dx < C\left(\sum_1^n \gamma_j\right)^2\Omega(n).$$

By the classical argument of Menchoff,[1] this would imply that the series $\sum \alpha_j\phi_j$ converges almost everywhere whenever $\sum \alpha_j^2\Omega(j) < \infty$. Hence, by the particular choice of the system $\{\phi_j\}$, a contradiction to Lemma I.

We shall now prove a third lemma, in which we shall adopt the following notations:

$\{\theta_k(t)\}_1^\infty$ is any orthonormal, uniformly bounded system in $(0,1)$, say $|\theta_k(t)| \leqq M$.

$\{c_k\}_1^\infty$ is any sequence of coefficients, and $s_m(t) = \sum_1^m c_k\theta_k(t)$.

$\{m_j\}_1^\infty$ is any increasing sequence of positive integers.

$\psi(t)$ is of the class L in $(0,1)$, and $\psi(t) \geqq 1$ (ψ may depend on n),

$\{\phi_j(x)\}_1^\infty$ is any orthonormal system in $(0,1)$.

$\{F_j(x)\}_1^n$ is any non-increasing sequence of characteristic functions of sets in $(0,1)$ (which may depend on n).

---

[1] We mean the argument of Menchoff's theorem used to prove that for all orthonormal systems $\{\phi_j\}_1^\infty$, $\int_0^1 \sup_{1\leqq p\leqq n} |S_p(x)|^2 dx < C(\sum_1^n \gamma_j^2)\log^2 n$ and deducing from this inequality that $\sum \gamma_j\phi_j$ converges almost everywhere whenever $\sum \gamma_j^2 \log^2 j < \infty$.

In order to apply this argument (see, e. g., Kaczmarz and Steinhaus [1]) it is neecssary to suppose that $\Omega(n)/\log n$ exceeds a constant. But this is irrelevant for the proof of Lemma II, since clearly we can replace, in (1), $\Omega(n)$ by a function increasing more rapidly.

$Q(x)$ is any function of the class L^2 in $(0,1)$ (which may depend on n), with $\int_0^1 Q^2(x)\,dx = 1$.

For the sake of simplicity, everything is real.

LEMMA III. *We have, with the preceding notations,*

$$\sum_{j=1}^{n} \Big(\int_0^1 QF_j\phi_j\,dx\Big)^2 c_{m_j}^2 \leqq M^2 \Big(\int_0^1 \psi\,dt\Big) \cdot \max_{1\leqq j\leqq n} \int_0^1 s_{m_j}(t)/\psi(t)\,dt.$$

Proof. The system of functions $p_j(x,t) = \phi_j(x)\theta_{m_j}(t)\psi(t)^{\frac{1}{2}}$ is orthogonal in the square $0 \leqq t \leqq 1$, $0 \leqq x \leqq 1$. Let

$$a_j = \int_0^1 \int_0^1 p_j^2(x,t)\,dx\,dt = \int_0^1 \theta_{m_j}^2(t)\psi(t)\,dt \leqq M^2 \int_0^1 \psi(t)\,dt, \tag{2}$$

and consider the function

$$P(x,t) = \sum_{k=1}^{m_n} c_k Q(x) f_k(x) \theta_k(t)/\psi(t)^{\frac{1}{2}},$$

where the $\{f_k(x)\}_1^{m_n}$ is a non-increasing sequence of m_n characteristic functions of sets in $(0,1)$ which will be determined in a moment.

One has, if $j \leqq n$, $\int_0^1 \int_0^1 P(x,t)p_j(x,t)\,dx\,dt = c_{m_j} \int_0^1 Qf_{m_j}\phi_j\,dx$, while the left-hand side vanishes if $j > n$. Hence Bessel's inequality gives

$$\sum_{j=1}^{n} \Big(\int_0^1 Qf_{m_j}\phi_j\,dx\Big)^2 c_{m_j}^2/a_j \leqq \int_0^1 \int_0^1 P^2(x,t)\,dx\,dt. \tag{3}$$

Now

$$P(x,t) = Q(x)/(\psi(t))^{\frac{1}{2}} \sum_{k=1}^{m_n} f_k(x) c_k \theta_k(t) = Q(x)/(\psi(t))^{\frac{1}{2}} \sum_{1}^{m_n} \Delta f_k \cdot s_k(t),$$

where $\Delta f_k = f_k(x) - f_{k+1}(x)$, $\Delta f_{m_n} = f_{m_n}$. Remembering the definition of $f_k(x)$, we have $P^2(x,t) = Q^2/\psi \sum_{1}^{m_n} \Delta f_k \cdot s_k^2(t)$ and

$$\int_0^1 \int_0^1 P^2(x,t)\,dx\,dt = \sum_{1}^{m_n} \Big(\int_0^1 Q^2 \Delta f_k\,dx\Big)\Big(\int_0^1 s_k^2(t)/\psi(t)\,dt\Big).$$

Take now

$$f_k = F_j \text{ for } m_{j-1} < k \leqq m_j \qquad (j = 1, 2, \cdots, n),\ (m_0 = 0),$$

and write $\Delta F_j = F_j - F_{j+1}$, $\Delta F_n = F_n$. Then

$$\int_0^1 \int_0^1 P^2(x,t)\,dx\,dt = \sum_{1}^{n} \Big(\int_0^1 Q^2 \Delta F_j\,dx\Big)\Big(\int_0^1 s_{m_j}^2(t)/\psi(t)\,dt\Big)$$
$$\leqq \max_{1\leqq j\leqq n} \int_0^1 s_{m_j}^2(t)/\psi(t)\,dt,$$

since $\int_0^1 Q^2 dx = 1$. Hence, by (3) and (2),

$$\sum_{j=1}^{n} \Bigl(\int_0^1 QF_j\phi_j dx\Bigr)^2 c_{m_j}{}^2 \leqq M^2 \Bigl(\int_0^1 \psi dt\Bigr) \max_{1\leqq j\leqq n} \int_0^1 s_{m_j}{}^2(t)/\psi(t)\, dt,$$

which proves the Lemma.

Proof of the theorem. The notations being the same as in the Lemma, take $\theta_k(t) = \cos(kt - \alpha_k)$, so that the series $\sum c_k\theta_k(t)$ becomes a trigonometric series (the fact that the interval of orthogonality is $(0, 2\pi)$ and that the system is not normalized is obviously irrelevant). We define now:

$$\psi(t) = \psi_n(t) = \sup_{1\leqq j\leqq n} |s_{m_j}(t)| + 1,$$

and we suppose that $|c_{m_j}| \geqq \delta > 0$. The lemma then gives, A being an absolute constant:

$$\sum_{j=1}^{n} \Bigl(\int_0^1 QF_j\phi_j dx\Bigr)^2 \leqq A/\delta^2 \Bigl(\int_0^{2\pi} \psi_n(t)\, dt\Bigr) \max_{1\leqq j\leqq n} \int_0^{2\pi} |s_{m_j}(t)|\, dt.$$

Now it is easily seen, since s_{m_j} is a partial sum of the polynomial s_{m_n}, that

$$(4) \qquad \int_0^{2\pi} \psi_n(t)\, dt \leqq C \log m_n \int_0^{2\pi} |s_{m_n}(t)|\, dt.$$

Hence

$$\sum_{j=1}^{n} \Bigl(\int_0^1 QF_j\phi_j dx\Bigr)^2 \leqq AC/\delta^2 \log m_n \cdot \max_{1\leqq j\leqq n} \Bigl\{\int_0^{2\pi} |s_{m_j}(t)|\, dt\Bigr\}^2.$$

Now suppose that $\log m_n = O(\log n)$. Then the second member of the last inequality is

$$O(\log n) \cdot \max_{1\leqq j\leqq n} \Bigl\{\int_0^{2\pi} |s_{m_j}(t)|\, dt\Bigr\}^2.$$

If we had $\lim \int_0^{2\pi} |s_{m_j}(t)|\, dt/(\log j)^{\frac{1}{2}} = 0$, the inequality would give

$$\sum_{j=1}^{n} \Bigl(\int_0^1 QF_j\phi_j dx\Bigr)^2 = o(\log^2 n),$$

the second member being independent of $\{\phi_j\}_1{}^\infty$, $\{F_j\}_1{}^n$ and Q. This is impossible, by Lemma II. Hence, if in the series $\sum_1^\infty c_k \cos(kt - \alpha_k)$, whose

m-th partial sum we denote by s_m, we have $|c_{m_j}| \geqq \delta > 0$, $\log m_j = O(\log j)$, then

$$\limsup \int_0^{2\pi} |s_{m_j}(t)|\, dt/(\log j)^{\frac{1}{2}} > 0. \tag{5}$$

This result includes obviously our theorem, which is thus proved.

Remark. Let us observe that if $\sum c_k \cos(kt - \alpha_k)$ is a Fourier-Stieltjes series, one gets a stronger result than (5): for then (4) can be replaced by $\int_0^{2\pi} \psi_n(t)\, dt \leqq B \log m_n$, B being a constant. We have then with the same hypothesis

$$\sum_{j=1}^{n} \Big(\int_0^1 QF_j\phi_j dx \Big)^2 = O(\log n) \max \int_0^{2\pi} |s_{m_j}(t)|\, dt,$$

which leads, instead of (5) to

$$\limsup \int_0^{2\pi} |s_{m_j}(t)|\, dt/\log j > 0. \tag{6}$$

This is, in particular, the case, if $\liminf \int_0^1 |s_m(t)|\, dt = O(1)$, since this relation implies that $\sum c_k \cos(kt - \alpha_k)$ is a Fourier-Stieltjes series.

Other consequences of Lemma III. Lemma III can be used to prove other results. We shall give one example, which generalizes a result obtained earlier (see Salem [2]).

By $\theta_k(t)$ we denote now generally, as in the Lemma, any orthonormal, uniformly bounded system in $(0, 1)$, with $|\theta_k(t)| \leqq M$. Writing as before $s_m = \sum_1^m c_k\theta_k$, we shall prove that if $|c_{m_j}| \log j \to \infty$ then $\sup_j |s_{m_j}(t)|$ cannot belong to L.

In fact, suppose $\sup_j |s_{m_j}(t)| \in L$, and choose, in the inequality of Lemma III, $\psi(t) = \sup_j |s_{m_j}(t)| + 1$. The inequality becomes

$$\sum_{j=1}^{n} \Big(\int_0^1 QF_j\phi_j dx \Big)^2 c_{m_j}{}^2 \leqq M^2 \Big\{ \int_0^1 \psi(t)\, dt \Big\}^2.$$

Now, if $|c_{m_j}| \log j \to \infty$, we can find a function $\omega(j) \to \infty$ such that $\omega(j)/\log j$ decreases and that $|c_{m_j}| \log j > \omega(j)$. Hence

$$(\omega(n)/\log n)^2 \sum_{j=1}^{n} \Big(\int_0^1 QF_j\phi_j dx \Big)^2 \leqq M^2 \Big(\int_0^1 \psi dt \Big)^2,$$

which again contradicts Lemma II, by a suitable choice of $\{\phi_j\}_1^\infty$, $\{F_j\}_1^n$ and Q.

Hence $\sup_j | s_{m_j}(t) | \varepsilon L$ implies $\liminf | c_{m_j} | \log j = O(1)$.

One deduces immediately (by considering a subsequence of $\{m_j\}$) that $\sup_j | s_{m_j}(t) | \varepsilon L$ implies $c_{m_j} \rightarrow 0$.

In particular, $\sup_m | s_m(t) | \varepsilon L$ implies $c_m \rightarrow 0$ for any orthonormal, uniformly bounded system. Direct proofs can be given of this result, a remark which I owe to A. Zygmund.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

BIBLIOGRAPHY.

[1] S. Kaczmarz and H. Steinhaus, *Orthogonalreihen*, Warszawa-Lwów, 1935.

[2] R. Salem, "A new proof of a theorem of Menchoff," *Duke Mathematical Journal*, vol. 8 (1941), pp. 269-272.

SOME LINEAR MINIMAX PROBLEMS OVER AN ORDERED FIELD.*

By R. J. LEVIT.

---

1. Introduction. Let S be a subset of an n-dimensional vector space over an ordered field F, fixed throughout the paper. If, as the vector $X = (X_1, \cdots, X_n)$ ranges over S, $\max_j |X_j|$ has a minimum value M^*, then M^* is called the *minimax of* S. Clearly, when M^* exists, it may be characterized as the unique element of F satisfying the two conditions,

$$\text{For every } X \varepsilon S \max_{j=1,\cdots,n} |X_j| \geqq M^*. \tag{1.1}$$
$$\text{There is a vector } X^* \varepsilon S \text{ such that } \max_{j=1,\cdots,n} |X^*_j| = M^*.$$

A vector X^* satisfying the second of these conditions is called a *minimax vector of* S. It may be regarded as least among the vectors of S in the sense that each component is at least as small in absolute value as some component of every other vector in S.

Given a consistent system of linear equations with coefficients in F,

$$\sum_{j=1}^{n} a_{ij}x_j = a_{i0} \qquad (i = 1, \cdots, m), \tag{1.2}$$

let G be a set of integers $g_1, \cdots, g_s$, $1 \leqq g_\nu \leqq n$, and denote by S^G the totality of vectors $(X_{g_1}, \cdots, X_{g_s})$ such that $x = (X_1, \cdots, X_n)$ is a solution of (1.2). The minimax M^G of S^G we call the *minimax of the system with respect to* $x_{g_1}, \cdots, x_{g_s}$. Thus, M^G is the minimum value of $\max_{j \varepsilon G} |X_j|$ as X ranges over all solutions of (1.2). A minimax vector X^G of S^G we call a *minimax solution with respect to* $x_{g_1}, \cdots, x_{g_s}$. Our main result is the determination of M^G and the characterization of the minimax solutions with respect to $x_{g_1}, \cdots, x_{g_s}$ for the system (1.2). If certain submatrices of the matrix of (1.2) are non-singular, there is a unique minimax solution X^G, and an explicit formula is given for it. These results are obtained in Section 4 utilizing the already known case of a single equation [1], which is discussed in Section 2, and the notion of a Plücker array, which is defined and developed in Section 3.

---

* Received July 7, 1953; revised December 15, 1954.

In Section 5 the results of Section 4 are applied to determine the minimax M^* of the range of a vector function $L(x_1, \cdots, x_q)$ having as components the linear functions,[1]

$$L_i(x_1, \cdots, x_q) = \sum_{j=1}^{q} b_{ij} x_j - b_{i0} \qquad (i = 1, \cdots, n), \tag{1.3}$$

as $x_1, \cdots, x_q$ range independently over F. In this case M^* is called *the minimax of the system of linear functions.* The vectors $X^* = (X^*_1, \cdots, X^*_q)$ such that $\max_i |L_i(X^*)| = M^*$, called *the minimax points of the system,* are characterized also.

In Section 6 the results of Section 4 and Section 5 are applied to some problems of best approximation in the Tchebycheff sense; namely determination of best approximate solutions to inconsistent systems of linear equations[2] and best approximation of arbitrary function on finite sets. The latter generalizes some results of De La Vallée Poussin [5].

2. The minimax solutions of a single equation.

We consider the equation,

$$\sum_{j=1}^{n} a_j x_j = a_0 \qquad (a_j \varepsilon F). \tag{2.1}$$

LEMMA 1.1. *If X is a solution of (2.1) and $a_j \neq 0$ for at least one integer j, $1 \leqq j \leqq n$, then $\max_{j=1,\cdots,n} |X_j| \geqq |\mu|$, where*

$$\mu = a_0 / \sum_{j=1}^{n} |a_j|. \tag{2.2}$$

Proof. The result follows immediately from

$$|a_0| \leqq \sum_{j=1}^{n} |a_j| \, |X_j| \leqq (\max |X_j|) \sum_{j=1}^{n} |a_j|. \tag{2.3}$$

THEOREM 1. *Denote the set $\{1, \cdots, n\}$ by N. If in (2.1) $a_j \neq 0$ for at least one $j \varepsilon N$, the solutions have the minimax $M^N = |\mu|$. Let D be the set of integers $j \varepsilon N$ for which $a_j = 0$. Then X^N is a minimax solution if and only if*

$$|X^N{}_j| \leqq |\mu| (j \varepsilon D); \qquad X^N{}_j = \mu \operatorname{sgn} a_j \quad (j \varepsilon N - D). \tag{2.4}$$

---

[1] This problem has been solved in the field of real numbers by M. Krein [3] using functional analysis and in the field of complex numbers by V. K. Ivanov [2].

[2] This problem has recently been treated in the field of complex numbers by E. Remez [4].

Proof. Consider first the case $a_0a_1\cdots a_n\neq 0$. By the lemma $|\mu|$ satisfies the first minimax condition (1.1). For a solution X of (2.1) equality holds between the first and second members of (2.3) if and only if the non-vanishing terms in the sum $\sum a_jX_j$ all have the same sign (which is then necessarily the sign of a_0) and between the second and third members if and only if the X_j all have the same absolute value. Hence, $\max|X_j|=|\mu|$ if and only if $\operatorname{sgn} X_j=\operatorname{sgn} a_0a_j=\operatorname{sgn}\mu a_j$ and $|X_j|=|\mu|$ for $j=1,\cdots,n$. Since direct substitution with the aid of (2.2) verifies that $x_j=\mu\operatorname{sgn} a_j$ is a solution of (2.1), $|\mu|$ also satisfies the second minimax condition (1.1) with the second expression (2.4) as the minimax solution, which is unique in this case. The proof for D empty is then completed by noting that, when $a_0=0$, the unique minimax solution is $X^N{}_j=0$, $j=1,\cdots,n$. Finally, when D is not empty, we apply the foregoing result to the system $\sum_{j\varepsilon N-D} a_jx_j=a_0$. Then, since $a_0/\sum_{j\varepsilon N-D}|a_j|=\mu$, the theorem is proved.

3. **Sign-compatibility of Plücker arrays.** Any skew-symmetric square array over F of rank two or less we call a *Plücker array* (P. a.). The existence proof for minimax solutions of (1.1) depends on the property of these arrays stated in Theorem 2 below, which involves the following notion. Two vectors,

$$(3.1)\qquad C_1=(c_{11},\cdots,c_{1s}),\qquad C_2=(c_{21},\cdots,c_{2s}),$$

over F are called *sign-compatible* (s. c.) if $\operatorname{sgn} c_{1j}c_{2j}c_{1k}c_{2k}\neq -1$ for $j,k=1,\cdots,s$. This relation is evidently symmetric and reflexive but is not transitive; e. g., of the vectors $(1,1,1)$, $(1,0,1)$, $(1,-1,1)$ the first and second are s. c. as are the second and third but not the first and third.

LEMMA 2.1. *Let the vectors* (3.1) *contain a pair of corresponding components* c_{1a} *and* c_{2a} *such that* $c_{1a}c_{2a}\neq 0$. *Then* C_1 *and* C_2 *are s. c. if and only if*

$$(3.2)\qquad \operatorname{sgn} c_{1j}c_{2j}=\operatorname{sgn} c_{1a}c_{2a}$$

whenever $c_{1j}c_{2j}\neq 0$.

Proof. If C_1 and C_2 are s. c. and $c_{1j}c_{2j}\neq 0$, $\operatorname{sgn} c_{1j}c_{2j}c_{1a}c_{2a}=1$, and (3.2) follows. Conversely, assume (3.2) holds whenever $c_{1j}c_{2j}\neq 0$. Then for arbitrary $j,k=1,\cdots,s$ either $c_{1j}c_{2j}c_{1k}c_{2k}=0$ or

$$\operatorname{sgn} c_{1j}c_{2j}=\operatorname{sgn} c_{1a}c_{2a}=\operatorname{sgn} c_{1k}c_{2k}\neq 0$$

so that $\operatorname{sgn} c_{1j}c_{2j}c_{1k}c_{2k}=1$. Hence, C_1 and C_2 are s. c.

LEMMA 2.2. *If vectors* C_p, C_q, C_r *are rows of a P.a.* $C=(c_{ij})$, *then*

$$c_{qr}C_p+c_{rp}C_q+c_{pq}C_r=0. \tag{3.3}$$

Proof. If

$$c_{pq}=c_{qr}=c_{rp}=0, \tag{3.4}$$

(3.3) holds trivially; hence, we assume (3.4) does not hold. Since the rank of C is not more than two, there are elements k_1, k_2, $k_3\,\varepsilon\,F$ not all zero such that

$$k_1c_{pj}+k_2c_{qj}+k_3c_{rj}=0, \qquad (j=1,\cdots,s) \tag{3.5}$$

where s is the order of C. For definiteness we assume $k_3\neq 0$. Then, writing (3.5) in the cases $j=p$ and $j=q$ and applying the skew-symmetry of C, we obtain respectively

$$-k_2c_{pq}+k_3c_{rp}=0, \qquad k_1c_{pq}-k_3c_{qr}=0.$$

These imply that $c_{pq}\neq 0$, since otherwise (3.4) would hold. Hence, $k_2=k_3c_{rp}/c_{pq}$, $k_1=k_3c_{qr}/c_{pq}$. Substituting these values into (3.5), we obtain (3.3).

LEMMA 2.3. *Two rows* C_p *and* C_q *of a P.a.* $C=(c_{ij})$ *are proportional if and only if* $c_{pq}=0$.

Proof. Let $c_{pj}=kc_{qj}$, $j=1,\cdots,s$, where s is the order of C and $k\,\varepsilon\,F$. Then in particular, $c_{pq}=kc_{qq}=0$, since C is skew-symmetric. Conversely, let $c_{pq}=0$. If $c_{pj}=0$ for $j=1,\cdots,s$, C_p and C_q are proportional. Otherwise, there is an element $c_{pr}\neq 0$; so $c_{rp}=-c_{pr}\neq 0$. By Lemma 2.2 $c_{qr}C_p+c_{rp}C_q=0$. Hence, C_p and C_q are proportional.

A P.a. is said to be reducible if two of its rows are proportional. By the lemma just proved a P.a. is irreducible (not reducible) if and only if all its zero elements lie in the main diagonal. Let $C=(c_{ij})$ be a reducible P.a. of order s with row C_b proportional to C_a.

$$c_{bj}=kc_{aj} \qquad (j=1,\cdots,s;\,k\,\varepsilon\,F). \tag{3.6}$$

The sub-array C' obtained from C by deleting the row and column of index b is called a *direct reduction of* C. An $(s-n)$-rowed minor array $C^{(n)}$ of C is called a *reduction of* C if there is a sequence $C', C'',\cdots, C^{(n-1)}$ of reducible sub-arrays of C each of which is a direct reduction of its immediate predecessor and such that C' is a direct reduction of C and $C^{(n)}$ of $C^{(n-1)}$. It is convenient to regard $C^{(0)}=C$ as itself a reduction of C. Clearly, a reduction of a P.a.

is again a P. a. We use the symbol $C^{(n)}_i$ to denote the row $(c_{ij_1}, \cdots, c_{ij_{s-n}})$ of $C^{(n)}$.

LEMMA 2.4. *Two rows $C^{(n)}_d$ and $C^{(n)}_e$ in a reduction $C^{(n)}$ of a reducible P.a. $C = (c_{ij})$ are proportional if and only if C_d and C_e are proportional and are s.c. if and only if C_d and C_e are s.c.*

Proof. It is sufficient to prove the lemma for $n = 1$; i.e., for a direct reduction C' of C. Let C'_d and C'_e be proportional. Then by Lemma 2.3 $c_{de} = 0$, and hence, C_d and C_e are proportional. Next, let C'_d and C'_e be s.c., and suppose that C' has been obtained from C by deleting the row and column of index b so that (3.6) holds for some integer a. Then,

$$\text{(3.7)} \qquad \operatorname{sgn} c_{dj}c_{ej}c_{dm}c_{em} \neq -1$$

holds for every $j, m = 1, \cdots, b-1, b+1, \cdots, s$. But

$$\operatorname{sgn} c_{dj}c_{ej}c_{db}c_{eb} = \operatorname{sgn} c_{dj}c_{ej}c_{bd}c_{be} = \operatorname{sgn} c_{dj}c_{ej}k^2c_{ad}c_{ae} = \operatorname{sgn} k^2c_{dj}c_{ej}c_{da}c_{ea} \neq -1$$

by (3.6), (3.7). That is, (3.7) holds for $m = b$ also so that C_d and C_e are s.c. The converses are trivial.

LEMMA 2.5. *If a P.a. $C = (c_{ij})$ contains a set of r rows no two of which are proportional, so does any reduction of C.*

Proof. It is sufficient to prove the lemma for a direct reduction C'. Let C' be obtained from C by deleting the row and column of index b so that (3.6) holds for some a. Let I be the set of indices of the r rows in question. If $b \notin I$, the conclusion is an immediate consequence of the preceding lemma. Otherwise, $a \notin I$ and the set $\{C'_i\}$ for $i \in I - \{b\} + \{a\}$ is composed of r rows of C'. By Lemma 2.4 no two of these rows are proportional except possibly C'_a and some row C'_d with $d \in I - \{b\}$. But, since C_b and C_d are not proportional, $c_{bd} \neq 0$ by Lemma 2.3. Hence, $c_{ad} \neq 0$ by (3.6), and accordingly by Lemma 2.3 C'_a and C'_d are not proportional either.

LEMMA 2.6. *Let (c_{ij}) be an irreducible P.a. of order $s \geqq 4$; define the vectors $V_{in} = (c_{i1}, c_{i2}, \cdots, c_{in})$, $n = 1, \cdots, s$, and let d, e, m be integers, $d \neq e$, $1 < d \leqq m < s$, $1 < e \leqq m$. Then if V_{1m} is s.c. with both V_{dm} and V_{em}, $V_{1,m+1}$ is s.c. with at least two of the vectors, $V_{d,m+1}$, $V_{e,m+1}$, $V_{m+1,m+1}$.*

Proof. If $V_{d,m+1}$ and $V_{e,m+1}$ are both s.c. with $V_{1,m+1}$, there is nothing to prove. Otherwise, we assume without loss of generality that $V_{d,m+1}$ is not s.c. with $V_{1,m+1}$. Then there is an integer a, $1 < a \leqq m$, such that $\operatorname{sgn} c_{1a}c_{da}c_{1,m+1}c_{d,m+1} = -1$, or $\operatorname{sgn} c_{1a}c_{da} = -\operatorname{sgn} c_{1,m+1}c_{d,m+1} \neq 0$. But, by

Lemma 2.1 for every integer $j \neq d$ such that $1 < j \leqq m$, $\operatorname{sgn} c_{1j}c_{dj} = \operatorname{sgn} c_{1a}c_{da}$, since $c_{ij} \neq 0$ for $j \neq i$ by Lemma 2.3. Hence,

$$(3.8) \qquad \operatorname{sgn} c_{1j}c_{dj} = -\operatorname{sgn} c_{1,m+1}c_{d,m+1} \neq 0 \qquad (1 < j \leqq m, j \neq d),$$

and in particular $\operatorname{sgn} c_{1e}c_{de} = -\operatorname{sgn} c_{1,m+1}c_{d,m+1} \neq 0$, so that because of the skew-symmetry

$$(3.9) \qquad \operatorname{sgn} c_{de}c_{1,m+1} = \operatorname{sgn} c_{e1}c_{d,m+1} \neq 0.$$

Similarly, if $V_{e,m+1}$ were not s.c. with $V_{1,m+1}$, $\operatorname{sgn} c_{ed}c_{1,m+1} = \operatorname{sgn} c_{d1}c_{e,m+1} \neq 0$, or $\operatorname{sgn} c_{de}c_{1,m+1} = \operatorname{sgn} c_{1d}c_{e,m+1} \neq 0$. This combined with (3.9) would imply that $c_{de}c_{1,m+1} + c_{e1}c_{d,m+1} + c_{1d}c_{e,m+1} \neq 0$ contrary to Lemma 2.2. Hence, $V_{e,m+1}$ is s.c. with $V_{1,m+1}$. Since by Lemma 2.2 $c_{m+1,1}c_{jd} + c_{1j}c_{m+1,d} + c_{j,m+1}c_{1d} = 0$ and, for $j \neq d$ such that $1 < j \leqq m$, $\operatorname{sgn} c_{1j}c_{m+1,d} = \operatorname{sgn} c_{m+1,1}c_{jd} \neq 0$ by (3.8), it follows that

$$\operatorname{sgn} c_{j,m+1}c_{1d} = -\operatorname{sgn} c_{1j}c_{m+1,d} \neq 0, \text{ or } \operatorname{sgn} c_{1j}c_{m+1,j} = \operatorname{sgn} c_{1d}c_{m+1,d} \neq 0$$

for $1 < j \leqq m$; i.e., whenever $c_{1j}c_{m+1,j} \neq 0$ by Lemma 2.3. Hence, by Lemma 2.1 $V_{m+1,m+1}$ is s.c. with $V_{1,m+1}$; and the lemma is proved.

THEOREM 2. *If a P.a. $C = (c_{ij})$ contains a set of three rows no two of which are proportional, then any given row C_a containing a non-zero element is s.c. with precisely two rows C_f and C_g not proportional to C_a or to each other. C_f and C_g are not s.c. with each other provided that C contains a row not proportional to any of the rows, C_a, C_f, C_g.*

Proof. We may suppose without loss of generality that C_a is C_1, since a P.a. remains a P.a. when the rows are renumbered in any order provided that the columns are also renumbered in that same order, and since sign-compatibility and proportionality of rows is independent of their numbering. First we establish the existence of the rows C_f and C_g when C is irreducible. In the notation of Lemma 2.6 V_{23} and V_{33} are s.c. with V_{13}, since $c_{ii} = 0$. When the order of C is $s = 3$, this means that C_2 and C_3 are s.c. with C_1. When $s \geqq 4$, it implies that at least two of the vectors, V_{24}, V_{34}, V_{44}, are s.c. with V_{14} by Lemma 2.6 with $m = 3$. Similarly, applying Lemma 2.6 with $m = 4, 5, \cdots, s-1$ successively, we find distinct integers f and g, $1 < f \leqq s$, $1 < g \leqq s$, such that V_{fs} and V_{gs} are s.c. with V_{1s}; i.e., C_f and C_g are s.c. with C_1. Now suppose C is reducible, containing n rows proportional to C_1 (besides C_1 itself). The minor array obtained by deleting these n rows and the corresponding columns is by Lemma 2.4 a reduction $C^{(n)}$ of C, in which

${C^{(n)}}_1$ is proportional to no other row. By Lemma 2.5 $C^{(n)}$ has an irreducible reduction $C^{(r)}$, $r \geqq n$, containing at least three rows, one of which is ${C^{(r)}}_1$. As already shown for irreducible arrays, $C^{(r)}$ contains two rows, ${C^{(r)}}_f$ and ${C^{(r)}}_g$, s.c. with ${C^{(r)}}_1$. Hence, by Lemma 2.4, C_f and C_g are s.c. with C_1, and no two of the three are proportional. Next, if there is a row C_h not proportional to any of the rows, C_a, C_f, C_g, then by Lemmas 2.3, 2.1 $\operatorname{sgn} c_{1g}c_{fg} = \operatorname{sgn} c_{1h}c_{fh}$, $\operatorname{sgn} c_{1f}c_{gf} = \operatorname{sgn} c_{1h}c_{gh}$, and hence, $\operatorname{sgn} c_{1g}c_{fh} = \operatorname{sgn} c_{1h}c_{fg} = -\operatorname{sgn} c_{1f}c_{gh}$. Therefore, $\operatorname{sgn} c_{1f}c_{hf}c_{1g}c_{hg} = \operatorname{sgn} c_{f1}c_{g1}c_{fh}c_{gh} = -1$. Thus, C_h cannot be s.c. with C_1, and neither can C_f be s.c. with C_g.

4. The minimax solutions of the general system.

If the system (1.2) is consistent and has rank $r > 0$, we can number the equations so that the subsystem,

$$\sum_{j=1}^{n} a_{ij}x_j = a_{i0} \qquad (i = 1, \cdots, r), \tag{4.1}$$

also has rank r. It is then clear that, with respect to any given subset of the unknowns, every minimax solution of (1.2) is a minimax solution of (4.1) and conversely. We, therefore, confine our attention to the system (4.1) with rank equal to the number of equations. Denote the matrix of (4.1) by A; i.e., $A = (a_{ij})$ is an $r \times n$ matrix of rank r. Denote the augmented matrix by $\bar{A}$, the $r \times s$ matrix composed of the columns of $\bar{A}$ with indices $j_1, \cdots, j_s$ in that order by $A_{j_1 \cdots j_s}$, and the determinant of the square matrix $A_{j_1 \cdots j_r}$ by $\alpha_{j_1 \cdots j_r}$. It is sometimes convenient to let J denote the ordered set $(j_1, \cdots, j_q)$, $1 \leqq q \leqq s \leqq n$, and write $A_{Jj_{q+1} \cdots j_s}$ for $A_{j_1 \cdots j_s}$ and similarly for the α's when $s = r$. If J is the null set, we understand $A_{Jj_1 \cdots j_s}$ to be the same as $A_{j_1 \cdots j_s}$. The determinants $\alpha_{j_1 \cdots j_r}$ play an important role in the minimax problem because they are essentially Plücker coordinates for the $(n-r)$-dimensional subspace of intersection of the r hyperplanes (4.1) in projective n-space over F, and they consequently characterize the totality of solutions independently of the particular system of equations through which it is defined. However, we shall derive the required properties of the α's directly without recourse to this geometrical interpretation. Evidently an interchange of subscripts merely changes the sign of $\alpha_{j_1 \cdots j_r}$; and, if two or more subscripts are alike, $\alpha_{j_1 \cdots j_r} = 0$. If $r > 1$ and $J = (j_1, \cdots, j_{r-1})$, we denote the $(r-1)$-rowed square submatrix obtained by deleting the row of index i from A_J by ${A^i}_J$ and its determinant by ${\alpha^i}_J$. We then have

$$\sum_{i=1}^{r} (-1)^{r+i} {\alpha^i}_J a_{ij} = \alpha_{Jj} \qquad (j = 0, \cdots, n). \tag{4.4}$$

LEMMA 3.1. *If* $(X_1,\cdots,X_n)$ *is a solution of the system* (4.1) *and* $J=(j_1,\cdots,j_{r-1})$,

$$\sum_{j=1}^{n}\alpha_{Jj}X_j=\alpha_{J0}. \tag{4.5}$$

Proof. By (4.4)

$$\sum_{j=1}^{n}\alpha_{Jj}X_j=\sum_{j=1}^{n}\sum_{i=1}^{r}(-1)^{r+i}\alpha^i{}_J a_{ij}X_j=\sum_{i=1}^{r}(-1)^{r+i}\alpha^i{}_J\sum_{j=1}^{n}a_{ij}X_j$$

$$=\sum_{i=1}^{r}(-1)^{r+i}\alpha^i{}_J a_{i0}=\alpha_{J0}.$$

By (4.4) if $J=(j_1,\cdots,j_{r-1})$, the last row of the matrix

$$\begin{pmatrix} a_{1k_1}\cdots a_{1k_{r+1}} \\ \cdots\cdots\cdots \\ a_{rk_1}\cdots a_{rk_{r+1}} \\ \alpha_{Jk_1}\cdots\alpha_{Jk_{r+1}} \end{pmatrix}$$

is a linear combination of the first r rows, and hence, the determinant vanishes. Expanding by elements of the last row, we obtain the bilinear relations.

$$\sum_{\nu=1}^{r+1}(-1)^{r+1+\nu}\alpha_{k_1\ldots k_{\nu-1}k_{\nu+1}\ldots k_{r+1}}\alpha_{j_1\ldots j_{r-1}k_\nu}=0. \tag{4.6}$$

LEMMA 3.2. *Let* $K=(k_1,\cdots,k_{r-2})$ *be a sequence of* $r-2$ *distinct integers,* $1\leqq k_\nu\leqq n$. *Then the array* (α_{Kij}) *for* $i,j=1,\cdots,n$ *is a Plücker array.*

Proof. Evidently, $\alpha_{Kij}=-\alpha_{Kji}$. Moreover, setting $j_h=k_h$ for $h=1,\cdots,r-2$, $j_{r-1}=j$, $k_{r-1}=d$, $k_r=e$, $k_{r+1}=f$ in (4.6), we obtain

$$\alpha_{Kef}\alpha_{Kdj}+\alpha_{Kfd}\alpha_{Kej}+\alpha_{Kde}\alpha_{Kfj}=0. \tag{4.7}$$

Thus, three arbitrarily chosen rows, $(\alpha_{Ki1},\cdots,\alpha_{Kin})$ for $i=d,e,f$, are linearly dependent.

The following familiar result is stated here for reference.

LEMMA 3.3. *If a matrix* B *has a non-singular* $r\times r$ *submatrix* C *and all* $(r+1)\times(r+1)$ *submatrices of* B *that contain* C *are singular, the rank of* B *is* r.

Let $J=(j_1,\cdots,j_{r-1})$ be a sequence of $r-1$ distinct integers, $1\leqq j_\nu\leqq n$. The number $w(J)$ of zero determinants in the sequence (α_{Jj}), $j=1,\cdots,n$, we call *the order of* J *with respect to* A. Evidently, $w(J)\geqq r-1$. If $w(J)<n$, we call J a *proper column index sequence* (p. c. i. s.) *for* A. Since

A is of rank $r > 0$, there is a determinant $\alpha_{j_1 \ldots j_r} \neq 0$. Hence, there is always a p. c. i. s. for A. (When $r = 1$, the (only) p. c. i. s. for A is the null set.)

LEMMA 3.4. *$J = (j_1, \cdots, j_{r-1})$ is a p. c. i. s. for A if and only if A_J has maximum rank $r - 1$.*

Proof. Let A_J have rank $r - 1$ so that, for some integer $g, A^g{}_J$ is non-singular. Let A' be a matrix obtained by permuting the columns of A so that A_J occupies the first $r - 1$ columns. The only $r \times r$ submatrices of A' that contain $A^g{}_J$ are the matrices A_{Jj} with $j \notin J$. They cannot all be singular, since by Lemma 3.3 the rank of A' would then be $r - 1$ instead of r. Hence, at least one of the determinants α_{Jj} does not vanish; so J is a p. c. i. s. for A' and hence, for A. Conversely, if $w(J) < n$, there is an integer f, $1 \leqq f \leqq n$, such that $\alpha_{Jf} \neq 0$ and hence, by (4.4) not all the minors $\alpha^i{}_J$ are zero. Therefore, the rank of A_J is $r - 1$.

LEMMA 3.5. *Let H be a sequence of $q < r$ integers $h_1, \cdots, h_q$, $1 \leqq h_\nu \leqq n$. If the matrix A_H has maximum rank q, then there is a p. c. i. s. for A containing H.*

Proof. Let A_H have rank q. If $q = r - 1$, H is itself a p. c. i. s. by Lemma 3.4. Next, let $q = r - 1 - s$ with $1 \leqq s < r - 1$. A_H contains a non-singular $q \times q$ submatrix Q. If every $(q + 1) \times (q + 1)$ submatrix of A containing Q is singular, then by Lemma 3.3 $r = q$ contrary to hypothesis. Hence, A has a non-singular $(q + 1) \times (q + 1)$ submatrix Q_1 containing Q. If $s = 1$, the order of Q_1 is $r - 1$; and by Lemma 3.4 the indices of the columns of A occurring in Q_1 constitute a p. c. i. s. containing H. If $s > 1$, we repeat the process obtaining a $(q + 2) \times (q + 2)$ submatrix Q_2 containing Q_1 and therefore Q. After s such steps we obtain a non-singular $(q + s) \times (q + s) = (r - 1) \times (r - 1)$ submatrix Q_s containing Q, and the indices of the columns of A occurring in Q_s constitute the desired p. c. i. s. containing H.

LEMMA 3.6. *Let the matrix A be such that the sequence $L = (1, 2, \cdots, r - 1)$ is a p. c. i. s. of order w and*

$$(4.8) \qquad \alpha_{Lj} = 0 \quad (j = 1, \cdots, w); \qquad \alpha_{Lj} \neq 0 \quad (j = w + 1, \cdots, n).$$

and let $J = (j_1, \cdots, j_{r-1})$ be a sequence of integers, $1 \leqq j_\nu \leqq w$. Then,

$$(4.9) \qquad \alpha_{Jk} = \theta_J \alpha_{Lk} \qquad (k = 1, \cdots, n).$$

where θ_J is an element of F independent of k; namely,

$$(4.10) \qquad \theta_J = \alpha_{Jn} / \alpha_{Ln}.$$

Proof. By Lemma 3.4 there is an integer g such that $A^g{}_L$ is non-singular. All the $r \times r$ submatrices of $A_{1\ldots w}$ containing $A^g{}_L$ are singular by (4.8). Hence, by Lemma 3.3 $\alpha_{Jk} = 0$, $k = 1, \cdots, w$; so that, setting $k_h = h$, $h = 1, \cdots, r-1$, $k_r = n$, $k_{r+1} = k$, (4.6) becomes $-\alpha_{Lk}\alpha_{Jn} + \alpha_{Ln}\alpha_{Jk} = 0$. Since L is a p. c. i. s., $\alpha_{Ln} \neq 0$ by (4.8) and the lemma follows.

LEMMA 3.7. *Let A be as in the preceding lemma with the further property that $\alpha^r{}_L \neq 0$, and let $K = (k_1, \cdots, k_{r-2})$ be a sequence of distinct integers, $1 \leqq k_\nu \leqq w$. Then, for each $j = 1, \cdots, n$*

$$\alpha_{Kjk} = \phi_j \alpha^r{}_{Kk} \qquad (k = 1, \cdots, w), \tag{4.11}$$

where ϕ_j is an element of F independent of k; namely,

$$\phi_j = -\alpha_{Lj}/\alpha^r{}_L. \tag{4.12}$$

Proof. By (4.4) and (4.8) we have

$$\alpha^r{}_L \alpha_{Kjk} = \begin{vmatrix} a_{1k_1} & \cdots a_{1k_{r-2}} & a_{1j} & a_{1k} \\ \cdot & \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & \cdot & \cdot \\ a_{r-1,k_1} & \cdots a_{r-1,k_{r-2}} & a_{r-1,j} & a_{r-1,k} \\ \alpha_{Lk_1} & \cdots \alpha_{Lk_{r-2}} & \alpha_{Lj} & \alpha_{Lk} \end{vmatrix} = -\alpha_{Lj}\alpha^r{}_{Kk} \quad (k = 1, \cdots, w)$$

LEMMA 3.8. *With the hypothesis and notation of Lemma 3.7, if K is a p. c. i. s. for $A^r{}_{1\ldots w}$, there is an integer e, $1 \leqq e \leqq w$, such that $\theta_{Ke} \neq 0$ and $\alpha_{Kej} \neq 0$ for $j = w+1, \cdots, n$.*

Proof. Since the order of K with respect to $A^r{}_{1\ldots w}$ is less than w, there is an integer e, $1 \leqq e \leqq w$, such that $\alpha^r{}_{Ke} \neq 0$. Hence, A_{Ke} has maximal rank $r-1$; so by Lemma 3.4 $(k_1, \cdots, k_{r-2}, e)$ is a p. c. i. s. for A. Thus, there is an integer q, $1 \leqq q \leqq n$, such that $\alpha_{Keq} \neq 0$. By Lemma 3.6 $\alpha_{Keq} = \theta_{Ke}\alpha_{Lq}$. Hence, $\theta_{Ke} \neq 0$, and therefore $\alpha_{Kej} = \theta_{Ke}\alpha_{Lj} \neq 0$ for $j = w+1, \cdots, n$ by (4.8).

LEMMA 3.9. *Under the hypothesis of Lemma 3.8 there are integers f and g, $w < f \leqq n$, $w < g \leqq n$, such that all the non-zero members in the set $\{\alpha_{Kgj}\alpha_{Lj}\}$, $j = w+1, \cdots, n$, have the sign of ϕ_g (defined by (4.12)) and all those in the set $\{\alpha_{Kfj}\alpha_{Lj}\}$ have the sign of $-\phi_f$.*

Proof. If $\alpha_{Kij} = 0$ for $i, j = w+1, \cdots, n$, the theorem is trivial. Hence we assume there are integers b and c, $w < b \leqq n$, $w < c \leqq n$, such that $\alpha_{Kbc} \neq 0$. By Lemma 3.8 there is an integer e, $1 \leqq e \leqq w$, such that $\alpha_{Kej} \neq 0$ for $j = w+1, \cdots, n$. Then the array (α_{Kij}) $(i, j = 1, \cdots, n)$, which by Lemma 3.2 is a P. a., contains three rows no two of which are proportional; namely, $R_e = (\alpha_{Ke1}, \cdots, \alpha_{Ken})$, $(\alpha_{Kb1}, \cdots, \alpha_{Kbn})$, $(\alpha_{Kc1}, \cdots, \alpha_{Kcn})$ by Lemma

2.3. Accordingly, by Theorem 2 there are two rows $R_f = (\alpha_{Kf1}, \cdots, \alpha_{Kfn})$ and $R_g = (\alpha_{Kg1}, \cdots, \alpha_{Kgn})$ s.c. with R_e and not proportional to R_e or to each other. By Lemma 2.3

$$\alpha_{Kef}\alpha_{Keg}\alpha_{Kfg} \neq 0; \tag{4.13}$$

and, since $\alpha_{Kek} = 0$ for $k = 1, \cdots, w$ by Lemma 3.7 and eq. (4.8), $f > w$ and $g > w$. Then for any integer j, $w < j \leqq n$, such that $\alpha_{Kfj} \neq 0$, $\operatorname{sgn} \alpha_{Kej}\alpha_{Kfj} = \operatorname{sgn} \alpha_{Keg}\alpha_{Kfg}$ by Lemma 2.1. But by Lemma 3.6 $\alpha_{Kej} = \theta_{Ke}\alpha_{Lj}$ and $\alpha_{Keg} = \theta_{Ke}\alpha_{Lg}$; and therefore, since $\theta_{Ke} \neq 0$ by Lemma 3.8,

$$\operatorname{sgn} \alpha_{Kfj}\alpha_{Lj} = \operatorname{sgn} \alpha_{Kfg}\alpha_{Lg}. \tag{4.14}$$

Likewise for any integer j, $w < j \leqq n$, such that $\alpha_{Kgj} \neq 0$,

$$\operatorname{sgn} \alpha_{Kgj}\alpha_{Lj} = \operatorname{sgn} \alpha_{Kgf}\alpha_{Lf}. \tag{4.15}$$

By Lemma 3.6 and eq. (4.12),

$$\operatorname{sgn}(\alpha_{Kef}\alpha_{Keg}/\alpha^r{}_L) = \operatorname{sgn}(\alpha_{Lf}\alpha_{Lg}/\alpha^r{}_L) = -\operatorname{sgn} \phi_f\alpha_{Lg} = -\operatorname{sgn} \phi_g\alpha_{Lf}. \tag{4.16}$$

By (4.13) one of the expressions, $\alpha_{Kef}\alpha_{Keg}\alpha_{Kfg}$ and $\alpha_{Kef}\alpha_{Keg}\alpha_{Kgf}$, is positive and the other negative. Hence, we may assign the subscripts f and g so that

$$\operatorname{sgn}(\alpha_{Kef}\alpha_{Keg}\alpha_{Kfg}/\alpha^r{}_L) = 1.$$

Therefore, by (4.16), $-\operatorname{sgn} \phi_f\alpha_{Lg}\alpha_{Kfg} = 1$, or $\operatorname{sgn} \alpha_{Kfg}\alpha_{Lg} = -\operatorname{sgn} \phi_f$ so that by (4.14),

$$\operatorname{sgn} \alpha_{Kfj}\alpha_{Lj} = -\operatorname{sgn} \phi_f \tag{4.17}$$

for all non-zero members of the set $\{\alpha_{Kfj}\alpha_{Lj}\}$, $j = w + 1, \cdots, n$. Again by (4.16) $-\operatorname{sgn} \phi_g\alpha_{Lf}\alpha_{Kfg} = 1$, or $\operatorname{sgn} \alpha_{Kgf}\alpha_{Lf} = \operatorname{sgn} \phi_g$ so that by (4.15),

$$\operatorname{sgn} \alpha_{Kgj}\alpha_{Lj} = \operatorname{sgn} \phi_g \tag{4.18}$$

for all non-zero members of the set $\{\alpha_{Kgj}\alpha_{Lj}\}$. (4.17) and (4.18) establish the lemma.

Let $J = (j_1, \cdots, j_{r-1})$ be a p.c.i.s. of order w with respect to A, and denote by $j_r, j_{r+1}, \cdots, j_n$ the integers $1, \cdots, n$ not in J in an order such that

$$\alpha_{Jj_\nu} = 0 \quad (\nu = 1, \cdots, w); \qquad \alpha_{Jj_\nu} \neq 0 \quad (\nu = w + 1, \cdots, n). \tag{4.19}$$

Hence, the system (4.1) (of rank r) has solutions with arbitrary values assigned to $x_{j_r}, x_{j_{r+1}}, \cdots, x_{j_{n-1}}$. In particular, if we define

$$\mu_J = \alpha_{J0}/\sum_{j=1}^{n} |\alpha_{Jj}|, \qquad \gamma_{Jj} = \mu_J \operatorname{sgn} \alpha_{Jj} \qquad (j = 1, \cdots, n) \tag{4.20}$$

($\sum_{j=1}^{n} | \alpha_{Jj} | \neq 0$ since J is a p.c.i.s. for A), there are elements $X_{j_1}, \cdots, X_{j_w}$ and X_{j_n} in F such that $x_j = \gamma_{Jj}$ for $j = j_{w+1}, \cdots, j_{n-1}$; $x_j = X_j$ for $j = j_1, \cdots, j_w, j_n$ is a solution of (4.1). Then by Lemma 3.1 and eqs. (4.20), (4.19),

$$\alpha_{J0} = \sum_{\nu=1}^{w} \alpha_{Jj_\nu} X_{j_\nu} + \sum_{\nu=w+1}^{n-1} \alpha_{Jj_\nu} \gamma_{Jj_\nu} + \alpha_{Jj_n} X_{j_n} = \mu_J \sum_{\nu=1}^{n-1} | \alpha_{Jj_\nu} | + \alpha_{Jj_n} X_{j_n}$$

$$= \alpha_{J0} + | \alpha_{Jj_n} | (X_{j_n} \operatorname{sgn} \alpha_{Jj_n} - \mu_J).$$

Hence, $X_{j_n} = \mu_J / \operatorname{sgn} \alpha_{Jj_n} = \gamma_{Jj_n}$. We conclude that the system,

$$\text{(4.21)} \qquad \sum_{\nu=1}^{w} a_{ij_\nu} x_{j_\nu} = a_{i0} - \sum_{\nu=w+1}^{n} a_{ij_\nu} \gamma_{Jj_\nu} \qquad (i = 1, \cdots, r),$$

in $x_{j_1}, \cdots, x_{j_w}$ has the solution $(X_{j_1}, \cdots, X_{j_w})$. The matrix $A_{j_1 \ldots j_w}$ of this system contains a non-singular $(r-1)$-rowed submatrix $A^g{}_J$ by Lemma 3.4; and by (4.19) all the $r \times r$ submatrices of $A_{j_1 \ldots j_w}$ containing $A^g{}_J$ are singular. Hence, by Lemma 3.3 the rank of $A_{j_1 \ldots j_w}$ is $r-1$. The set of equations (4.21) is called the *derived system of* (4.1) *with respect to* J. We have just proved

LEMMA 3.10. *The derived system of* (4.1) *with respect to a p.c.i.s. for A is consistent and has rank $r-1$.*

We now define

$$\text{(4.22)} \qquad a'_{i0} = a_{i0} - \sum_{j=w+1}^{n} a_{ij} \gamma_{Lj}; \; \alpha'_{K0} = \begin{vmatrix} a_{1k_1} & \cdots & a_{1k_{r-2}} & a'_{10} \\ \cdot & \cdot & \cdot & \cdot \\ a_{r-1,k_1} & \cdots & a_{r-1,k_{r-2}} & a'_{r-1,0} \end{vmatrix};$$

$$\mu'_K = \alpha'_{K0} / \sum_{k=1}^{w} | \alpha^r{}_{Kk} |$$

where K and L have the meanings assigned in Lemmas 3.6 and 3.8.

LEMMA 3.11. *Under the hypothesis of Lemma 3.8, there is an integer h, $w < h \leqq n$, such that either* $| \mu_{Kh} | = | \mu'_K |$ or $| \mu_L | \gtrless | \mu_{Kh} | \gtrless | \mu'_K |$.

Proof. By Lemma 3.10 the derived system of (4.1) with respect to the p.c.i.s. $L = (1, \cdots, r-1)$ has a solution $(X_1, \cdots, X_w)$, so that with the notation (4.22)

$$\text{(4.23)} \qquad \sum_{k=1}^{w} a_{ik} X_k = a'_{i0} \qquad (i = 1, \cdots, r);$$

and, assuming a suitable numbering of equations (4.1), the first $r-1$ equations of the derived system form a system of rank $r-1$. Applying Lemma 3.1 to this latter system,

$$(4.24)\qquad \sum_{k=1}^{w} \alpha^{r}{}_{Kk} X_k = \alpha'_{K0}.$$

Then by Lemma 3.7, eq. (4.24), (4.22), and Lemma 3.7 respectively,

$$(4.25)\qquad \sum_{k=1}^{w} \alpha_{Kjk} X_k = \phi_j \sum_{k=1}^{w} \alpha^{r}{}_{Kk} X_k = \phi_j \alpha'_{K0} = \phi_j \mu'_K \sum_{k=1}^{w} | \alpha^{r}{}_{Kk} |$$

$$= \mu'_K (\operatorname{sgn} \phi_j) \sum_{k=1}^{w} | \alpha_{Kjk} | \qquad (j = w+1, \cdots, n).$$

But by (4.23), (4.22), $(X_1, \cdots, X_w, \gamma_{L.w+1}, \cdots, \gamma_{Ln})$ is a solution of (4.1), and by Lemma 3.8 $Kj = (k_1, \cdots, k_{r-2}, j)$ is a p.c.i.s. for A when $w < j \leqq n$. Accordingly,

$$(4.26)\qquad \mu_{Kj} \sum_{k=1}^{n} | \alpha_{Kjk} | = \alpha_{Kj0} = \sum_{k=1}^{w} \alpha_{Kjk} X_k + \sum_{k=w+1}^{n} \alpha_{Kjk} \gamma_{Lk}$$

$$= \mu'_K (\operatorname{sgn} \phi_j) \sum_{k=1}^{w} | \alpha_{Kjk} | + \sum_{k=w+1}^{n} \alpha_{Kjk} \gamma_{Lk}$$

for $j = w+1, \cdots, n$ by (4.20), Lemma 3.1, (4.25) respectively. Let f and g be the integers whose existence was established in Lemma 3.9; define $h = f$ if $\mu_L \mu'_K < 0$ and $h = g$ otherwise, so that all non-zero terms in the last member of (4.26) have the same sign for $j = h$. Then we obtain from (4.26)

$$| \mu_{Kh} | \sum_{k=1}^{n} | \alpha_{Khk} | = | \mu'_K | \sum_{k=1}^{w} | \alpha_{Khk} | + | \mu_L | \sum_{k=w+1}^{n} | \alpha_{Khk} | ;$$

and hence

$$(| \mu_{Kh} | - | \mu'_K |) \sum_{k=1}^{w} | \alpha_{Khk} | = (| \mu_L | - | \mu_{Kh} |) \sum_{k=w+1}^{n} | \alpha_{Khk} | ,$$

from which the lemma follows at once since $\sum_{k=1}^{w} | \alpha_{Khk} | \neq 0$ by Lemma 3.8.

We now define the *residual submatrix of a system of linear equations* (1.2) *with respect to* $x_{g_1}, \cdots, x_{g_s}$ as the matrix obtained by deleting the columns of index $g_1, \cdots, g_s$ from the matrix of the system, and proceed to state the principal theorem.

THEOREM 3. *Denote the set* $\{1, \cdots, n\}$ *by* N; *let* $G = \{g_1, \cdots, g_s\}$ *be a set of* s *integers of* N, *and write* $H = N - G$. *Let the matrix* A *of the system* (4.1) *have rank* r *equal to the number of equations. Also, if* $s < n$, *let* A_H, *the residual submatrix of* (4.1) *with respect to* $x_{g_1}, \cdots, x_{g_s}$, *have maximum rank. Then the system has the minimax* M^G *with respect to these* *unknowns given by the following*: (a) *For* $1 \leqq s \leqq n-r$, $M^G = 0$ *and the minimax solutions with respect to* $x_{g_1}, \cdots, x_{g_s}$ *are the vectors* X^G *such that*

$$(4.27) \qquad X^G{}_j = 0 \quad (j \varepsilon G); \qquad \sum_{j \varepsilon H} a_{ij} X^G{}_j = a_{i0} \qquad (i = 1, \cdots, r).$$

(b) *For* $n-r < s \leqq n$, $M^G = \max_J |\mu_J|$, where μ_J *is defined by* (4.20) *and* J *ranges over the set* Γ_H *consisting*[4] *of all the p.c.i.s. for* A *that contain* H. *Moreover, let* E *be a p.c.i.s. in* Γ_H *such that* $|\mu_E| = \max_{J \varepsilon \Gamma_H} |\mu_J|$ *and let* D *be the set of integers* $j \varepsilon N$ *such that*

$$(4.28) \qquad \alpha_{Ej} = 0 \quad (j \varepsilon D); \qquad \alpha_{Ej} \neq 0 \qquad (j \varepsilon N - D).$$

Then the minimax solutions with respect to $x_{g_1}, \cdots, x_{g_s}$ *are the vectors* X^G *such that*

$$(4.29) \qquad X^G{}_j = \gamma_{Ej} \qquad (j \varepsilon N - D)$$

where γ_{Ej} *is defined in* (4.20) *and the remaining components satisfy the equations,*

$$(4.30) \qquad \sum_{j \varepsilon D} a_{ij} X^G{}_j = a_{i0} - \sum_{j \varepsilon N-D} a_{ij} \gamma_{Ej} \qquad (i = 1, \cdots, r),$$

and the inequalities,

$$(4.31) \qquad |X^G{}_j| \leqq |\mu_E| \qquad (j \varepsilon D - H).$$

Proof. First let $1 \leqq s \leqq n-r$. Then, since the matrix A_H of the coefficients of the unknowns x_j with $j \varepsilon H$ has the same rank r as A and $\bar{A}$, (4.1) has solutions with arbitrary values assigned to the remaining unknowns $x_{g_1}, \cdots, x_{g_s}$. In particular, there is a solution X^G satisfying the conditions (4.27), and the theorem follows at once for this case from the minimax criterion (1.1). For $n-r < s \leqq n$ the proof is by induction on r. For For $r = 1$ (in which case $s = n$) the present theorem reduces to Theorem 1. Next let $r = R > 1$, $n - R < s \leqq n$, and assume that the theorem holds for $r = R - 1$, $n - (R-1) < s \leqq n$. For any solution X of (4.1)

$$\sum_{j \varepsilon N-E} \alpha_{Ej} X_j = \alpha_{E0}$$

by Lemma 3.1, since $\alpha_{Ej} = 0$ for $j \varepsilon E$. Applying Lemma 1.1 to this equation and noting that $G = N - H \supseteq N - E$, we find that

$$\max_{j \varepsilon G} |X_j| \geqq \max_{j \varepsilon N-E} |X_j| \geqq |\mu_E|.$$

---

[4] Γ_H is non-empty by Lemma 3.5. If $s = n$, H is empty and Γ_H is understood to consist of all p.c.i.s. for A.

Then by the minimax criterion (1.1), if there is a solution X^G of (4.1) such that

$$\max_{j\varepsilon G} |X^G{}_j| = |\mu_E|, \tag{4.32}$$

then $M^G = |\mu_E|$, and X^G is a minimax solution of (4.1) with respect to $x_{g_1}, \cdots, x_{g_s}$. Moreover, X^G satisfies the conditions (4.29)-(4.31). To verify this last statement note that by Lemma 3.1, (4.28), $x_j = X^G{}_j$ is a solution of the equation, $\sum_{j\varepsilon N-D} \alpha_{Ej} x_j = \alpha_{E0}$; and at the same time $\max_{j\varepsilon N-D} |X^G{}_j| \leqq |\mu_E|$ by (4.32), since $G \supseteqq N-D$. Therefore by Theorem 1 and (4.28), it is a minimax solution of this equation and has components $X^G{}_j$ for $j \varepsilon N-D$ given by (4.29). Then, since X^G satisfies (4.1), the remaining components satisfy (4.30); while (4.31) follows from (4.32). Conversely, since $|\gamma_{Ej}| = |\mu_E|$, any vector X^G satisfying (4.29)-(4.31) is a solution of (4.1) for which (4.32) holds. It remains to establish the existence of a vector X^G such that the components $X^G{}_j$ for $j \varepsilon D$ satisfy (4.30) and (4.31). For this purpose it is convenient to assume the unknowns numbered so that the subscripts in H are $1, 2, \cdots, n-s$ (unless $s = n$, in which case H is empty); the subscripts in $E-H$ are $n-s+1, n-s+2, \cdots, R-1$ (unless $s = n-R+1$, in which case $E-H$ is empty); and the subscripts in $D-E$ (if any) are $R, R+1, \cdots, w$, where w is the order of E with respect to A. Then $E = L = (1, \cdots, R-1)$, and the equations (4.21) for the derived system of (4.1) with respect to E become

$$\sum_{j=1}^{w} a_{ij} x_j = a_{i0} - \sum_{j=w+1}^{n} a_{ij} \gamma_{Lj} \qquad (i = 1, \cdots, R). \tag{4.33}$$

By Lemma 3.10, assuming a suitable numbering of the equations (4.1), the first $R-1$ equations (4.33) form a system S' of rank $R-1$. For $s < n$, if A_H has maximum rank, then the residual matrix $A^R{}_H$ of S' with respect to $x_{n-s+1}, x_{n-s+2}, \cdots, x_w$ also has maximum rank, since the R-th row of A_H is linearly dependent on the other rows. Hence, if $s > n-(R-1)$, then by the induction hypothesis S' has with respect to these unknowns the minimax, $M' = \max_K |\mu'_K|$ as $K = (k_1, \cdots, k_{R-2})$ ranges over all p.c.i.s. for $A^R{}_{1\ldots w}$ containing $(1, \cdots, n-s)$. With the specified numbering of the unknowns (4.28) reduces to (4.8), and all the hypotheses of Lemmas 3.6-3.8 and hence, also 3.11 apply to A and each set K with $r = R$. If for any such K, $|\mu'_K| > |\mu_L|$, then by Lemma 3.11 there would be an integer h, $w < h \leqq n$, such that $|\mu_{Kh}| > |\mu_L|$. But, inasmuch as (Kh) is a p.c.i.s. for A by Lemma 3.8 and contains H since K does, we have $|\mu_{Kh}| \leqq \max_{J\varepsilon\Gamma_H} |\mu_J| = |\mu_E|$

$= |\mu_L|$. Hence, $M' = \max\limits_K |\mu'_K| \leqq |\mu_L|$ for $s > n - (R-1)$. This holds for $s = n - (R-1)$ also, since in that case S' has the minimax $M' = 0$ with respect to $x_{n-s+1}, \cdots, x_w$, as we find by applying the present theorem for $s \leqq n - r$, in which case it has already been proved. Now, if $(X^G{}_1, \cdots, X^G{}_w)$ is any solution of S', then by Lemma 3.10 it satisfies all the equations (4.33) and hence, also with the specified numbering the conditions (4.30). If in addition it is a minimax solution of S' with respect to $x_{n-s+1}, \cdots, x_w$, then $|X^G{}_j| \leqq M' \leqq |\mu_L|$ for $j = n-s+1, \cdots, w$, and with the specified numbering this is equivalent to (4.31). That is, we obtain the required vector X^G by choosing $(X^G{}_1, \cdots, X^G{}_w)$ as a minimax solution of S' with respect to $x_{n-s+1}, \cdots, x_w$. Thus, the theorem is proved for $r = R$, $n - R < s \leqq n$, and the induction is complete.

We remark that in determining $\max\limits_J |\mu_J|$ it is sufficient to consider sequences $J = (j_1, \cdots, j_{r-1})$ with $j_1 < j_2 < \cdots < j_{r-1}$, since the value of $|\mu_J|$ is independent of the order of the indices in J.

Corollary 3. *For $n - r < s \leqq n$, if $D = E = (e_1, \cdots, e_{r-1})$; i.e., if the determinants α_{Ej} vanish only for $j \in E$, there is a single minimax solution X^G, the components $X^G{}_j$ for $j \in E$ being uniquely determined by the equations* (4.30) *alone. In the notation of* (4.22) *with E in place of L and the equations* (4.1) *numbered so that the forms $\sum\limits_{j \in E} a_{ij}x_j$ $(i = 1, \cdots, r-1)$ are linearly independent, this solution is explicitly given by the equations,*

$$X^G{}_{e_k} = \alpha'{}_{e_1 \ldots e_{k-1} 0 e_{k+1} \ldots e_{r-1}} / \alpha^r{}_E \quad (k = 1, \cdots, r-1); \qquad X^G{}_j = \gamma_{Ej} \quad (j \in N - E).$$

There remains the case that the residual matrix in question does not have maximum rank. This can be treated by means of the following theorem.

Theorem 4. *Let the matrix A of the system* (4.1) *have rank r equal to the number of equations. In the notation of Theorem* 3, *if $s < n$ and A_H, the residual matrix with respect to $x_{g_1}, \cdots, x_{g_s}$, has rank $t < \min(r, n-s)$, let T be a sequence of t integers out of H such that A_T has maximum rank t. Then* (4.1) *has the same minimax with respect to $x_{g_1}, \cdots, x_{g_s}$ as the system,*

$$\sum_{j \in T+G} a_{ij}u_j = a_{i0} \qquad (i = 1, \cdots, r), \tag{4.34}$$

has with respect to $u_{g_1}, \cdots, u_{g_s}$ (and the latter minimax can be determined by Theorem 3). *For each $k \in H - T$ there are elements c_{kj} such that*

$$\sum_{j \in T} a_{ij}c_{kj} = a_{ik} \qquad (i = 1, \cdots, r). \tag{4.35}$$

Each minimax solution X^G *of* (4.1) *can be derived from a corresponding minimax solution* U^G *of* (4.34) *by assigning arbitrary values to the components* $X^G{}_j$ *with* $j \varepsilon H - T$ *and determining the remaining components by means of the equations,*

$$(4.36) \qquad X^G{}_j = U^G{}_j \quad (j \varepsilon G), \qquad X^G{}_j = U^G{}_j - \sum_{k \varepsilon H-T} c_{kj} X^G{}_k \quad (j \varepsilon T).$$

Proof. For any integer $k \varepsilon H - T$ the augmented matrix of (4.35) (considered as a system of equations in the unknowns c_{kj}), being the same as a submatrix of A_H except possibly for a permutation of columns, has the same rank t as the matrix A_T of the system. Consequently there are elements c_{kj} satisfying (4.35). If X is a solution of (4.1) and elements U_j are given by

$$(4.37) \qquad U_j = X_j \quad (j \varepsilon G), \qquad U_j = X_j + \sum_{k \varepsilon H-T} c_{kj} X_k \quad (j \varepsilon T),$$

then $u_j = U_j$ is a solution of (4.34), since by (4.37), (4.35)

$$\sum_{j \varepsilon T} a_{ij} U_j = \sum_{j \varepsilon T} a_{ij} X_j + \sum_{j \varepsilon T} \sum_{k \varepsilon H-T} a_{ij} c_{kj} X_k = \sum_{j \varepsilon T} a_{ij} X_j + \sum_{k \varepsilon H-T} a_{ik} X_k = \sum_{j \varepsilon H} a_{ij} X_j.$$

Conversely, if $u_j = U_j$ is a solution of (4.34) and elements X_j for $j \varepsilon H - T$ are chosen arbitrarily, then elements X_j for $j \varepsilon T + G$ are uniquely determined by (4.37) and the resulting vector $(X_1, \cdots, X_n)$ is a solution of (4.1). In particular, if $u_j = U^G{}_j$ is a minimax solution of (4.34), then the vector X^G given by (4.36) with $X^G{}_j$ arbitrary for $j \varepsilon H - T$ is a solution of (4.1). Moreover, it is a minimax solution; for, if X is any other solution of (4.1) and the elements U_J are determined by (4.37), then, as noted above, $u_j = U_j$ is a solution of (4.34) and hence,

$$\max_j |X_j| = \max_j |U_j| \geqq \max_j |U^G{}_j| = |\max_j |X^G{}_j| \qquad (j \varepsilon G).$$

Conversely, X^G is a minimax solution of (4.1) only if the elements $U^G{}_j$ determined by (4.36) are the components of a minimax solution U^G of (4.34). Finally, the matrix of the system (4.34) has rank r equal to the number of equations, since it is obtained from A by deleting columns linearly dependent on those occurring in A_T; and the residual matrix with respect to $u_{g_1}, \cdots, u_{g_s}$ is A_T, which has maximum rank by hypothesis. Accordingly, (4.34) does indeed have a minimax solution U^G by Theorem 3. Hence, (4.1) has a minimax solution X^G and for both systems $M^G = \max_j |X^G{}_j| = \max_j |U^G{}_j|$ $(j \varepsilon G)$.

5. The minimax of a system of linear functions. The minimax and minimax points of a system of linear functions over F,

$$(5.1)\qquad L_i(x_1,\cdots,x_q)=\sum_{j=1}^{q}b_{ij}x_j-b_{i0}\qquad (i=1,\cdots,r),$$

can be determined by applying the theorems of the preceding section to the system of linear equations,

$$(5.2)\qquad \sum_{j=1}^{q}b_{ij}x_j-\sum_{j=q+1}^{q+r}\delta_{q+i,j}x_j=b_{i0}\qquad (i=1,\cdots,r),$$

where δ_{hj} is the unit element of F when $h=j$ and is the zero element otherwise. All solutions of (5.2) are given by

$$(5.3)\qquad x_j=X_j\quad (j=1,\cdots,q),\qquad x_j=L_{j-q}(X_1,\cdots,X_q)\quad (j=q+1,\cdots,q+r),$$

in which $X_1,\cdots,X_q$ are arbitrary elements of F. Consequently, the minimax $M^*(L)$ of (5.1) is the same as the minimax of (5.2) with respect to $x_{q+1},\cdots,x_{q+r}$. Moreover, $(X^*_1,\cdots,X^*_q)$ is a minimax point and $(L^*_1,\cdots,L^*_r)$ a minimax vector of (5.1) if and only if $(X^*_1,\cdots,X^*_q, L^*_1,\cdots,L^*_r)$ is a minimax solution of (5.2) with respect to $x_{q+1},\cdots,x_{q+r}$.

For convenient expression of the results we introduce some additional notations. If the rows of index $i_1,\cdots,i_s$ are deleted from a matrix B, we denote the resulting matrix by $B^{i_1\cdots i_s}$. If the superscripts are all distinct and the latter matrix is square, $\beta^{i_1\cdots i_s}$ is its determinant and otherwise is zero. Let B be $r\times q$ and, if $q<r$, let $I=(i_1,\cdots,i_{r-q-1})$ be a sequence of $r-q-1$ distinct integers, $1\leqq i_\nu\leqq r$. The number $d(I)$ of zeros in the sequence $(\beta^{I1},\cdots,\beta^{Ir})$ we call the *degree of* I *with respect to* B. Clearly, $d(I)\geqq r-q-1$. If $d(I)<r$, we call I a *proper row index* sequence (p. r. i. s.) *for* B.

THEOREM 5. *Let the matrix* $B=(b_{ij})$ $(i=1,\cdots,r;j=1,\cdots,q)$ *of the system of linear functions* (5.1) *have maximum rank, and denote by* B_0 *the matrix obtained from* B *by adjoining the elements* b_{i0} *as the* $(q+1)$*-th column. Then the system* (5.1) *has a minimax* $M^*(L)$. *If* $q\geqq r$, $M^*(L)=0$ *and the minimax points* X^* *are those for which* $L_i(X^*_1,\cdots,X^*_q)=0$ $(i=1,\cdots,r)$. *If* $q<r$, $M^*(L)=\max|\mu^I|$, *where* I *ranges over all p. r. i. s. for* B *and* μ^I *is defined by*

$$\mu^I=\beta_0{}^I/\sum_{i=1}^{r}|\beta^{Ii}|.$$

Let V *be a p. r. i. s. such that* $|\mu^V|=\max_I|\mu^I|$; *let* R *denote the set* $\{1,\cdots,r\}$, U *the set of integers* $i\varepsilon R$ *such that* $\beta^{Vi}=0$, *and* $\varepsilon(V,i)$ *the number of* *integers in V exceeding i. Then the minimax points X^* are those which satisfy the conditions,*

$$L_i(X^*_1,\cdots,X^*_q)=(-1)^{r-1+i+\varepsilon(V,i)}\mu^V\operatorname{sgn}\beta^{Vi}\qquad(i\,\varepsilon\,R-U). \tag{5.4}$$

$$|L_i(X^*_1,\cdots,X^*_q)|\leqq|\mu^V|\qquad(i\,\varepsilon\,U). \tag{5.5}$$

Proof. For $q\geqq r$ the theorem is immediate. To establish it in the case $q<r$ we proceed to determine the minimax M^G of (5.2) with respect to $x_{q+1},\cdot\cdot,x_{q+r}$, using the notation of Section 4 with $a_{ij}=b_{ij}$ for $j=0,1,\cdots,q$ and $a_{ij}=-\delta_{q+i,j}$ for $j=q+1,\cdots,q+r$. The rank of the matrix A of (5.2) turns out to be r, the same as the number of equations; and the residual matrix with respect to these unknowns is B, which has maximum rank by hypothesis. Thus, M^G can be determined by the case $n-r<s\leqq n$ of Theorem 3. Noting that for a sequence of integers $I=(i_1,\cdots,i_{r-q-1})$ with $1\leqq i_1<i_2<\cdots<i_{r-q-1}\leqq r$

$$\alpha_{1\ldots q,q+i_1,\ldots,q+i_{r-q}}=(-1)^p\beta^{Ii_{r-q}},$$

$$p=\tfrac{1}{2}(r-q)(r+q+3)+i_1+\cdots+i_{r-q}+\varepsilon\,(I,i_{r-q})$$

$$\alpha_{1\ldots q,q+i_1,\ldots,q+i_{r-q-1},0}=(-1)^{p_0}\beta_0{}^I,$$

$$p_0=\tfrac{1}{2}(r-q-1)(r+q+2)+i_1+\cdots+i_{r-q-1}$$

and hence, that J is a p.c.i.s. for A containing $H=\{1,\cdots,q\}$ if and only if $J=\{1,\cdots,q,q+i_1,\cdots,q+i_{r-q-1}\}$ and I is a p.r.i.s. for B, we have $\mu_J=(-1)^{p_0}\mu^I$ for $J\,\varepsilon\,\Gamma_H$. Consequently, M^G and hence, also $M^*(L)$ is $\max_I|\mu^I|$ as I ranges over the p.r.i.s. for B. Moreover, letting

$$V=(v_1,\cdots,v_{r-q-1})\text{ and }E=\{1,\cdots,q,q+v_1,\cdots,q+v_{r-q-1}\},$$

we have $|\mu_E|=\max_{J\varepsilon\Gamma_H}|\mu_J|$. Hence, in the notation of Theorem 3, $q+i\,\varepsilon\,D-H$ is equivalent to $i\,\varepsilon\,U$, and $q+i\,\varepsilon\,N-D$ is equivalent to $i\,\varepsilon\,R-U$, since $\alpha_{E,q+i}=0$ if and only if $\beta^{Vi}=0$. Then for any minimax vector X^* of (5.1) and any minimax solution X^G of (5.2) it follows from (4.29) that

$$\begin{aligned}L_i(X^*_1,\cdots,X^*_q)&=X^G{}_{q+i}=\gamma_{Eq+i}=\mu_E\operatorname{sgn}\alpha_{Eq+i}\\&=(-1)^{r-1+i+\varepsilon(V,i)}\mu^V\operatorname{sgn}\beta^{Vi}\qquad(i\,\varepsilon\,R-U),\end{aligned}$$

and from (4.31) that

$$|L_i(X^*_1,\cdots,X^*_q)|=|X^G{}_{q+i}|\leqq|\mu_E|=|\mu^V|\qquad(i\,\varepsilon\,U).$$

Since (4.30) is then satisfied automatically, the conditions (5.4), (5.5) characterize the minimax points X^* completely.

COROLLARY 5. *If $U = V$, i. e. if none of the determinants β^{Vi} $(i \notin V)$ vanish, the system* (5.1) *has a unique minimax point X^*, namely, the (in this case) unique solution of* (5.4).

To determine the minimax of (5.1) when the matrix B does not have maximum rank, we employ Theorem 4 and obtain

THEOREM 6. *Let the rank of B be $t < \min(q, r)$, and let T be a set of t integers out of $H = \{1, \cdots, q\}$ such that the submatrix B_T has maximum rank t. Then* (5.1) *has the same minimax as the system,*

$$\sum_{j \varepsilon T} b_{ij} u_j - b_{i0} \qquad (i = 1, \cdots, r) \tag{5.6}$$

with matrix of mikimum rank. For each $k \varepsilon H - T$ there are elements c_{kj} such that

$$\sum_{j \varepsilon T} b_{ij} c_{kj} = b_{ik} \qquad (i = 1, \cdots, r).$$

Each minimax point X^* of (5.1) *can be derived from a corresponding minimax point U^* of* (5.6) *by assigning arbitrary values to the components X^*_j with $j \varepsilon H - T$ and determining the remaining components by means of the equations,*

$$X^*_j = U^*_j - \sum_{k \varepsilon H - T} c_{kj} X^*_k \qquad (j \varepsilon T).$$

6. Problems in best approximation. The foregoing results can be applied to certain problems of best approximation in the Tchebycheff sense. For example, if as an alternative to the usual least squares criterion we define a *best approximate solution* x^* of an inconsistent system of linear equations,

$$\sum_{j=1}^{q} b_{ij} x_j = b_{i0} \qquad (i = 1, \cdots, r),$$

as a set of values for x which minimizes the largest of the "errors" $|\sum_j b_{ij} x_j - b_{i0}|$ $(i = 1, \cdots, r)$, then the vectors x^* are precisely the minimax points of the system of linear functions (5.1) and hence, are determined by Theorems 5 and 6.

Next we consider the problem of best approximation to a function at a finite number of points. Let S_p be a set of p points $x_1, \cdots, x_p$ in a completely arbitrary space, and let Φ be the class of all single-valued functions defined on S_p with range in the ordered field F. Given a subclass G of Φ and a function $\phi(x) \varepsilon \Phi$, we define a *G-function of best approximation to ϕ on S_p* as a function $g^*(x) \varepsilon G$ which minimizes $\max_{x \varepsilon S_p} |g(x) - \phi(x)|$ as g ranges over G. If $g_1(x), \cdots, g_q(x)$ are given functions in Φ and we take G as the class of functions expressible in the form $\sum_{j=1}^{q} c_j g_j(x)$ with $c_j \varepsilon F$, then[5] $g^*(x) = \sum_{j=1}^{q} c^*_j g_j(x)$, where $(c^*_1, \cdots, c^*_q)$ is a minimax point of the system of linear functions,

$$L_i(c_1, \cdots, c_q) = \sum_{j=1}^{q} c_j g_j(x_i) - \phi(x_i) \qquad (i = 1, \cdots, p),$$

and so again is determined by Theorems 5 and 6. The particular cases $g_j(x) = x^j$ and $g_{2k}(x) = \cos kx$, $g_{2k+1}(x) = \sin kx$ yield the results of De La Vallée Poussin [5] on best approximation on sets of p real numbers by polynomials of degree $n = q - 1 \leqq p - 2$ and by trigonometric sums of order $n = (q-1)/2 \leqq (p-2)/2$ respectively, if we note that in these cases none of the determinants β^{Ii} vanish so that Corollary 5 is applicable and there is a unique minimax point.

Finally, we determine the G-function of best approximation to ϕ on S_p when G is the class of all functions $g \varepsilon \Phi$ whose values on S_p are subject to a set of linear conditions,

$$(6.1) \qquad \sum_{j=1}^{p} c_{ij} g(x_j) = c_{i0} \qquad (i = 1, \cdots, r),$$

in which the c_{ij} are given elements of F. An important instance is the class of all functions in Φ whose n-th divided differences on S_p have preassigned values, $[x_i, x_{i+1}, \cdots, x_{i+n}] = c_{i0}$ for $i = 1, \cdots, p - n$. Then (6.1) holds with $p = r + n$,

$$c_{ij} = \prod_k (x_j - x_k)^{-1} (k = i, i+1, \cdots, j-1, j+1, \cdots, i+n)$$

for $j = i, i+1, \cdots, i+n$ and $c_{ij} = 0$ otherwise,[6] provided S_p is a subset of F. For this choice of the class of approximating functions the values $g^*(x)$ on S_p are given by

$$g^*(x_j) = Y^*_j + \phi(x_j) \qquad (j = 1, \cdots, p),$$

where $y = Y^*$ is any minimax solution of the system of linear equations,

$$\sum_{j=1}^{p} c_{ij} y_j = c_{i0} - \sum_{j=1}^{p} c_{ij} \phi(x_j) \qquad (i = 1, \cdots, r),$$

and can therefore be obtained by means of Theorems 3 and 4.

UNIVERSITY OF GEORGIA.

---

[5] This problem has recently been treated for $q = p - 1$ in the field of complex numbers by T. S. Motzkin and J. L. Walsh [8].

[6] See for example [7].

REFERENCES.

[1] Hans Rademacher and I. J. Schoenberg, "Helly's theorem on convex domains and Tchebycheff's approximation problem," *Canadian Journal of Mathematics*, vol. 2 (1950), pp. 251-252.

[2] V. K. Ivanov, "The problem of the minimax of a system of linear functions," *Mat. Sbornik N. S.* 28(70) (1951), pp. 685-706 (Russian); English abstract in *Mathematics Review*, vol. 13 (1952), p. 119.

[3] N. Ahiezer and M. Krein, *On Some Problems of the Theory of Moments*, Harkov, 1938 (Russian), pp. 171-199.

[4] E. Ya. Remez, "On Chebyshev approximations in a complex region," *Doklady Akad. Nauk SSSR* (N. S.), vol. 77 (1951), pp. 965-968 (Russian); English abstract in *Mathematics Review*, vol. 13 (1952), p. 99.

[5] C. de la Vallée Poussin, *Leçons sur l'approximation des fonctions d'une variable réelle*, Paris, 1919, Chapters VI and VII.

[6] M. Bôcher, *Introduction to Higher Algebra*, New York, 1907, p. 54.

[7] T. Fort, *Finite Differences*, Oxford, 1948, p. 10.

[8] T. S. Motzkin and J. L. Walsh, "On the derivative of a polynomial and Chebyshev approximation," *Proceedings of the American Mathematical Society*, vol. 4 (1953), pp. 76-87.

CONVERGENCE IN AREA OF INTEGRAL MEANS.*[1]

By CASPER GOFFMAN[2]

The integral means $f_n(x,y) = n^2 \int_0^{1/n} \int_0^{1/n} f(x+u, y+v)du\,dv$, $n=1, 2,\cdots$, of a function $f(x,y)$ have been used as a smoothing approximation to $f(x,y)$ in a variety of applications. Their use in surface area theory seems to have been first made by T. Radó, [1], who showed, for continuous $f(x,y)$, that the Lebesgue areas, $A(f_n)$, of the surfaces given by $f_n(x,y)$, $n=1,2,\cdots$, converge to the Lebesgue area $A(f)$ of the surface given by $f(x,y)$.

For summable $f(x,y)$, the Lebesgue area is in general meaningless and, in previous papers, L. Cesari [2] and the author [3] have introduced, in different ways, generalized Lebesgue areas, $\Phi_C(f)$ and $\Phi(f)$, which were proved in [3] to coincide for every summable $f(x,y)$. Our main purpose is to show that the result of Radó mentioned above concerning the $f_n(x,y)$ holds for every summable $f(x,y)$ provided the Lebesgue area is replaced by the generalized Lebesgue area. Indeed, for every summable $f(x,y)$, the integral means $f_n(x,y)$ are continuous so that their Lebesgue areas $A(f_n)$ do exist. What we show is that the sequence $\{A(f_n)\}$ either converges, or diverges to $+\infty$; i.e., it converges on the extended closed interval $[0,+\infty]$, and the limit coincides with the generalized Lebesgue area:

$$\lim_{n\to\infty} A(f_n) = \Phi(f) = \Phi_C(f).$$

It follows that the limit above can itself be interpreted as a definition of "the area of the surface defined by $f(x,y)$."

The area $H(f)$ of the paper [3], where we proved that $\Phi_C = \Phi = H$, is denoted by $\chi(f)$ in the present paper. The area $\Phi(f)$ of the present paper was not considered in [3] and we prove that $\Phi(f) = \chi(f)$.

For convenience, we say that a function $f(x,y)$ has *property R* if the sequence $\{A(f_n)\}$ of Lebesgue areas of its integral means converges on the interval $[0,+\infty]$. We thus show that every summable $f(x,y)$ has property

---

* Received November 3, 1954.

[1] Research supported by a grant from the National Science Foundation.

[2] We are grateful to the referee for his thoughtful criticism of the first version of this paper.

R. The property R is defined similarly for representations of parametric curves and surfaces. We shall show also that every $f: x_i(t)$, $i = 1, 2$, $0 \leqq t \leqq 1$, $x_i(t)$ summable, has property R. For parametric surfaces, the situation is difficult, as may be expected. For this case, we extend a result of C. B. Morrey [4], [5], [6], but shall omit the proofs since they are essentially the same as for the continuous case.

We first show that property R is general for curves; i. e., if $x(t)$, $y(t)$ are summable functions on the closed interval $[0, 1]$, with integral means $x_n(t) = n \int_0^{1/n} x(t+u)du$ and $y_n(t) = n \int_0^{1/n} y(t+u)du$, $n = 1, 2, \cdots$, then the sequence of lengths of the curves given by $x_n(t)$, $y_n(t)$ converges on the closed interval $[0, +\infty]$.

The generalized variation $\phi(x)$ of a summable function $x(t)$, on $[0, 1]$, was defined in [3]. One of several equivalent definitions is the following:

For any finite set $S: 0 < t_0 < t_1 < \cdots < t_n < 1$, let the norm, $\nu(S)$, be the largest of the numbers $t_0, t_1 - t_0, \cdots, 1 - t_n$. Let

$$\phi(x; S) = \sum_{i=1}^{n} | x(t_i) - x(t_{i-1}) |.$$

Now, let E be the set of points of approximate continuity of $x(t)$, and let $\phi(x) = \sup \phi(x; S)$ for all $S \subset E$. The following result is a simple consequence of this definition:

If $\phi(x) < +\infty$, then for every $\epsilon > 0$ there is a $\delta > 0$ such that $S \subset E$ and $\nu(S) < \delta$ implies $\phi(x; S) > \phi(x) - \epsilon$. If $\phi(x) = +\infty$ then for every $M > 0$ there is a $\delta > 0$ such that $S \subset E$ and $\nu(S) < \delta$ implies $\phi(x; S) > M$.

The function $\phi(x)$ agrees with the variation $V(x)$ for continuous $x(t)$ but not necessarily for discontinuous $x(t)$. However, it is always true that $\phi(x) \leqq V(x)$.

The integral means

$$x_n(t) = n \int_0^{1/n} x(t+u)du \text{ and } y_n(t) = n \int_0^{1/n} y(t+u)du$$

are defined on the interval $[0, 1 - 1/n]$, are absolutely continuous, and so are differentiable almost everywhere. It follows by the fundamental theorem of Lebesgue integration that the derivatives are given almost everywhere by

$$x_n'(t) = n\{x(t + 1/n) - x(t)\} \text{ and } y_n'(t) = n\{y(t + 1/n) - y(t)\}$$

so that, since $x_n(t)$ and $y_n(t)$ are absolutely continuous,

$$\phi(x_n) = V(x_n) = \int_0^{1-1/n} | x_n'(t) | \, dt = n \int_0^{1-1/n} | x(t + 1/n) - x(t) | \, dt$$

and

$$\phi(y_n) = n \int_0^{1-1/n} | y(t+1/n) - y(t) | \, dt.$$

We show that the sequences $\{\phi(x_n)\}$ and $\{\phi(y_n)\}$ converge on the closed interval $[0, +\infty]$. It is only necessary to show this for $\{\phi(x_n)\}$.

Let us first prove that $\phi(x_n) \leqq \phi(x)$. Indeed, we have

$$\phi(x_n) = n \int_0^{1-1/n} | x(t+1/n) - x(t) | \, dt$$
$$= n \int_0^{1/n} \sum_{i=1}^{n-1} | x(t+i/n) - x(t+(i-1)/n) \, dt.$$

For almost all t, $0 \leqq t \leqq 1/n$, the points $t, t+1/n, \cdots, t+(n-1)/n$ are all points of approximate continuity for $x(t)$, and hence, the sum under the last integral does not exceed $\phi(x)$. Thus:

$$\phi(x_n) \leqq n \int_0^{1/n} \phi(x) \, dt = \phi(x).$$

Suppose $\phi(x) < \infty$. Let $\epsilon > 0$. By the above remark, there is a $\delta > 0$ such that $S \subset E$ and $\nu(S) < \delta$ imply $\phi(x; S) > \phi(x) - \epsilon$. But, for all $n > 1/\delta$, and all $t \, \varepsilon \, [0, 1/n]$, the set $t, t+1/n, \cdots, t+(n-1)/n$ has norm less than δ and, for almost all t, it is a subset of E, so that

$$\phi(x) - \epsilon < \sum_{i=1}^{n-1} | x(t+i/n) - x(t+(i-1)/n) |.$$

Hence,

$$\phi(x) - \epsilon < n \int_0^{1/n} \sum_{i=1}^{n-1} | x(t+i/n) - x(t+(i-1)/n) | \, dt$$
$$= n \int_0^{1-1/n} | x(t+1/n) - x(t) | \, dt = \phi(x_n).$$

Thus, for every $n > 1/\delta$, we have $\phi(x) - \epsilon < \phi(x_n) \leqq \phi(x)$, so that $\{\phi(x_n)\}$ converges to $\phi(x)$. If $\phi(x) = +\infty$, a similar argument shows that $\{\phi(x_n)\}$ converges to $+\infty$.

Now, for each n, the pair $x_n(t)$, $y_n(t)$ represents a continuous curve C_n whose length is given by

$$L(C_n) = \int_0^{1-1/n} [x_n'(t)^2 + y_n'(t)^2]^{\frac{1}{2}} \, dt$$
$$= n \int_0^{1-1/n} [\{x(t+1/n) - x(t)\}^2 + \{y(t+1/n) - y(t)\}^2]^{\frac{1}{2}} \, dt$$
$$= n \int_0^{1/n} \sum_{i=1}^{n-1} [\{x(t+i/n) - x(t+(i-1)/n)\}^2$$
$$+ \{y(t+i/n) - y(t+(i-1)/n)\}^2]^{\frac{1}{2}} \, dt.$$

If, now, E is the set of points at which both $x(t)$ and $y(t)$ are approximately continuous, and f is the pair $x(t)$, $y(t)$, we may let

$$\psi(f;S)=\sum_{i=1}^{n}[\{x(t_i)-x(t_{i-1})\}^2+\{y(t_i)-y(t_{i-1})\}^2]^{\frac{1}{2}}$$

for each $S: 0<t_0<t_1<\cdots<t_n<1$ and then let $\psi(f)=\sup\psi(f;S)$ for all $S\subset E$. As for $\phi(f)$, we again have: If $\psi(f)<+\infty$, then for every $\epsilon>0$ there is a $\delta>0$ such that if $\nu(S)<\delta$, and $S\subset E$, then $\psi(f;S)>\psi(f)-\epsilon$; if $\psi(f)=+\infty$, then for every $M>0$ there is a $\delta>0$ such that if $\nu(S)>\delta$, and $S\subset E$, then $\psi(f;S)>M$.

It then follows, just as for $\phi(x)$, that if $\psi(f)<+\infty$ then for every $\epsilon>0$ there is an N such that $n>N$ implies $\psi(f)-\epsilon<L(C_n)<\psi(f)$, and if $\psi(f)=+\infty$, for every $M>0$ there is an N such that $n>N$ implies $L(C_n)>M$. We may now state:

THEOREM 1. *If $x(t)$, $y(t)$ are summable functions defined on $[0,1]$, and C_n is the curve corresponding to the integral means $x_n(t)=n\int_0^{1/n}x(t+u)\,du$, $y_n(t)=n\int_0^{1/n}y(t+u)\,du$, $n=1,2,\cdots$, then the sequence $\{L(C_n)\}$ of lengths converges on $[0,+\infty]$. In other words, every summable $x(t)$, $y(t)$ has property R.*

Let $f(x,y)$ be a summable function defined on the unit square $I=[0,1]\times[0,1]$. For every natural number n, the integral mean $f_n(x,y)=n^2\int_0^{1/n}\int_0^{1/n}f(x+u,y+v)\,du\,dv$ is defined on the square

$$I_n=[0,1-1/n]\times[0,1-1/n].$$

It is continuous, and is absolutely continuous in each variable for almost all values of the other variable; moreover, it is well known, and easy to prove, that its partial derivatives are given almost everywhere by the expressions

$$\partial f_n(x,y)/\partial x=n^2\int_0^{1/n}\{f(x+1/n,y+v)-f(x,y+v)\}\,dv \tag{1}$$

$$\partial f_n(x,y)/\partial y=n^2\int_0^{1/n}\{f(x+u,y+1/n)-f(x+u,y)\}\,du. \tag{2}$$

For continuous $f(x,y)$, [5], the derivatives $\partial f_n/\partial x$, $\partial f_n/\partial y$ are given everywhere by (1) and (2) and it is obvious that they are continuous.

It follows easily that $\partial f_n/\partial x$ and $\partial f_n/\partial y$ are summable. For:

$$\int_0^{1-1/n}\int_0^{1-1/n} | \partial f_n/\partial x | \, dx\, dy$$

$$= \int_0^{1-1/n}\int_0^{1-1/n} \{ n^2 \int_0^{1/n} | f(x+1/n, y+v) - f(x, y+v) | \, dv \} \, dx\, dy$$

$$\leqq n^2 \int_0^{1/n} dv \{ \int_0^{1-1/n}\int_0^{1-1/n} [| f(x+1/n, y+v) | + | f(x, y+v)] dx\, dy \}$$

$$\leqq 2n^2 \int_0^{1/n} dv \int_0^1\int_0^1 | f(x,y) | \, dx\, dy = 2n \int_0^1\int_0^1 | f(x,y) | \, dx\, dy,$$

and the result follows since $f(x,y)$ is summable.

Moreover, since $f_n(x,y)$ is absoluately continuous in each variable for all values of the other variable, it is ACT and so:

If $f(x,y)$ is summable, the surfaces given by the integral means $f_n(x,y)$, $n = 1, 2, \cdots$, all have finite Lebesgue areas given by

$$(3) \quad A(f_n) = \int_0^{1-1/n}\int_0^{1-1/n} [1 + \{ n^2 \int_0^{1/n} (f(x+1/n, y+v) - f(x, y+v)) dv \}^2$$
$$+ \{ n^2 \int_0^{1/n} (f(x+u, y+1/n) - f(x+u, y)) du \}^2]^{\frac{1}{2}} \, dx\, dy.$$

We shall show that the sequence (3) converges on the closed interval $[0, +\infty]$. It seems worth noting that the technique used in proving Theorem 1 now yields an inequality instead of an equality. Thus, a separate argument is needed to prove the opposite inequality. Since $f(x,y)$ is measurable, there are subsets E_x and E_y of $[0,1]$, each of measure 1, such that for every $x \varepsilon E_x$, $y \varepsilon E_y$, $f(x,y)$ is approximately continuous almost everywhere in the other variable.

Just as we did in the case of functions of one variable, we first prove our result for variation functions $\Phi_1(f)$ and $\Phi_2(f)$, which we now define. Let $\phi(f; x; c, d)$ be the generalized variation of $f(x,y)$ for x fixed and y varying on the interval $[c,d]$, and let $\phi(f; y; a, b)$ have similar meaning. We note that these functions are essentially additive; i.e., if $c < c' < d$ and $f(x,y)$ is approximately continuous in y at the point (x, c'), then

$$\phi(f; x; c, c') + \phi(f; x; c', d) = \phi(f; x; c, d).$$

We let

$$\Phi_1(f; a, b; c, d) = \int_a^b \phi(f; x; c, d) dx \quad \text{and} \quad \Phi_2(f; a, b; c, d) = \int_c^d \phi(f; y; a, b) dy.$$

In particular, we write,

$$\Phi_1(f) = \Phi_1(f; 0, 1; 0, 1) \quad \text{and} \quad \Phi_2(f) = \Phi_2(f; 0, 1; 0, 1).$$

For every a, b, c, d, with $0 \leqq a < b \leqq 1$, $0 \leqq c < d \leqq 1$, the functions $\Phi(f; x; c, d)$ and $\phi(f; y; a, b)$ are defined almost everywhere, have values in the interval $[0, +\infty]$, and are measurable, and so the above integrals have meaning.

The function $f(x, y)$ is of generalized bounded variation if $\Phi_1(f)$ and $\Phi_2(f)$ are both finite; otherwise, it is of generalized unbounded variation. For continuous $f(x, y)$, our definition agrees with that of bounded variation in the sense of Tonelli, and in the general case, agrees with Cesari's definition [2] of a function of type gBVT (or of generalized variation in the sense of Tonelli).

Since, for every n, the integral mean $f_n(x, y)$ is ACT, it follows that

$$\Phi_2(f_n) = \int_0^{1-1/n} \int_0^{1-1/n} | \partial f_n / \partial y | \, dx \, dy$$

$$\text{and } \Phi_2(f_n) = \int_0^{1-1/n} \int_0^{1-1/n} | \partial f_n / \partial x | \, dx \, dy.$$

We retain the convention of convergence on the closed interval $[0, +\infty]$, and we prove

THEOREM 2. *For every summable $f(x, y)$, the integral means $f_n(x, y)$ satisfy*

$$\lim_{n \to \infty} \Phi_1(f_n) = \Phi_1(f) \text{ and } \lim_{n \to \infty} \Phi_2(f_n) = \Phi_2(f).$$

Proof. We prove the theorem for $\Phi_2(f)$. Since

$$\partial f_n(x, y) / \partial x = n^2 \int_0^{1/n} \{ f(x + 1/n, y + v) - f(x, y + v) \} dv$$

almost everywhere, it follows that

$$\Phi_2(f_n) = \int_0^{1-1/n} \int_0^{1-1/n} n^2 \left| \int_0^{1/n} \{ f(x + 1/n, y + v) - f(x, y + v) \} dv \right| dx \, dy$$

$$\leqq \int_0^{1-1/n} \int_0^{1-1/n} n^2 \int_0^{1/n} | f(x + 1/n, y + v) - f(x, y + v) | dv \, dx \, dy$$

$$= \int_0^{1-1/n} n^2 \, dy \int_0^{1/n} \int_0^{1/n} \sum_{k=1}^{n-1} | f(x + k/n, y + v) - f(x + (k-1)/n, y + v) | \, dx \, d$$

$$= n^2 \int_0^{1/n} \int_0^{1/n} dx \, dv \{ \int_0^{1-1/n} \sum_{k=1}^{n-1} | f(x + k/n, y + v) - f(x + (k-1)/n, y + v) |$$

$$= n^2 \int_0^{1/n} dv \int_v^{v+1-1/n} dt \int_0^{1/n} \sum_{k=1}^{n} | f(x + k/n, t) - f(x + (k-1)/n, t) | \, dx.$$

For every v and for all $t \varepsilon E_y \cdot [v, v + 1 - 1/n]$, that is for almost all

$t \,\varepsilon\, [v, v+1-1/n]$, we have: for almost all x, the points $(x,t), (x+1/n, t)$, $\cdots, (x+(n-1/n, t)$ are all points of approximate continuity for $f(x,y)$ with respect to x, and hence:

$$\sum_{k=1}^{n} | f(x+k/n, t) - f(x+(k-1)/n, t) | \leqq \phi(f, t).$$

Then for the same v and t we have

$$\int_0^{1/n} \sum_{k=1}^{n} | f(x+k/n, t) - f(x+(k-1)/n, t) | \leqq n^{-1}\phi(f, t),$$

and finally

$$\Phi_2(f_n) \leqq n^2 \int_0^{1/n} dv \int_v^{v+1-1/n} dt \cdot n^{-1}\phi(f, t)$$

$$\leqq n^2 \int_0^{1} dt \int_0^{1/n} dv \cdot n^{-1}\phi(f, t) = \int_0^1 \phi(f, t) dt = \Phi_2(f).$$

We have thus shown that $\Phi_2(f_n) \leqq \Phi_2(f)$, for every n, so that

$$\limsup_{n\to\infty} \Phi_2(f_n) \leqq \Phi_2(f).$$

It remains for us to show that $\liminf_{n\to\infty} \Phi_2(f_n) \geqq \Phi_2(f)$. For this, we need the well known fact that $f_n(x,y)$ converges almost everywhere to $f(x,y)$ so that it converges almost everywhere in each variable for almost all values of the other variable.

We suppose that $\Phi_2(f) < +\infty$ and let $\epsilon > 0$. We then have $\phi(f;y) < +\infty$ for almost all values of y, and for every such y there is $b(y) < 1$ for which

$$\phi(f;y;0,b(y)) > \phi(f;y) - \epsilon/2,$$

where $\phi(f;y) = \phi(f;y;0,1)$. It follows that there is a $b < 1$ such that,

$$\int_0^1 \phi(f;y;0,b)\,dy > \Phi_2(f) - \epsilon/2.$$

Moreover, there is $a < 1$ such that

$$\int_0^a \phi(f;y;0,b)\,dy > \Phi_2(f) - \epsilon/2.$$

Let n_1 be a positive integer such that $n_1^{-1} < \min(1-a, 1-b)$. It is easy to see, and proved in [3], that the generalized variation $\phi(f)$ of a function $f(x)$ of one variable is lower semicontinuous with respect to almost everywhere convergence. It follows that, for almost all values of y, we have

$$\liminf_{n\to\infty} \phi(f_n;y;0,b) \geqq \phi(f;y;0,b).$$

Hence, for almost every y, there is an $n(y)$ such that $n > n(y)$ implies

$$\phi(f;y;0,b) > \phi(f;y;0,b) - \epsilon/2.$$

There is then a single N, which may be taken so that $N > n_1$, such that for $n > N$,

$$\Phi_2(f_n) \geqq \int_0^{1-1/n} \phi(f_n;y;0,b)dy \geqq \int_0^{1-1/n} \phi(f;y;0,b)dy - \epsilon/2$$

$$\geqq \int_0^a \phi(f;y;0,b)dy - \epsilon/2 \geqq \Phi_2(f) - \epsilon.$$

This completes the proof for the case $\Phi_2(f) < +\infty$. The case $\Phi_2(f) = +\infty$ provides no additional difficulties, so that Theorem 2 is proved.

Now, for every n, we have the inequalities

$$\Phi_1(f_n) + \Phi_2(f_n) + 1 \geqq A(f_n) \geqq \max[\Phi_1(f_n), \Phi_2(f_n)],$$

where $A(f_n)$ is the Lebesgue area of the integral mean $f_n(x,y)$. It follows that, if $f(x,y)$ is of generalized unbounded variation, the sequence $\{A(f_n)\}$ converges to $+\infty$, so that $f(x,y)$ has the property R.

To show that $\{A(f_n)\}$ converges when $f(x,y)$ is of generalized bounded variation, we consider the generalized Geöcze expression $\Psi(f)$. A subdivision ρ of I into rectangles $R_1, R_2, \cdots, R_m$ by means of lines parallel to the coordinate axes is *admissible* if $f(x,y)$ is approximately continuous in x for almost all y on all boundary lines of the rectangles of the subdivision parallel to the y axis, and in y for almost all x on all boundary lines parallel to the x axis. The meaning of the notation $\Phi_1(f;R_i)$, $\Phi_2(f;R_i)$, $i = 1, 2, \cdots, m$, should be clear, and it follows from the essential additivity of the generalized variation function $\phi(f;a,b)$ that

$$\Phi_1(f) = \sum_{i=1}^{m} \Phi_1(f;R_i) \text{ and } \Phi_2(f) = \sum_{i=1}^{m} \Phi_2(f;R_i).$$

We then consider the expression

$$\Psi(f;\rho) = \sum_{i=1}^{m} \Psi(f;R_i) = \sum_{i=1}^{m} [\{\Phi_1(f;R_i)\}^2 + \{\Phi_2(f;R_i)\}^2 + |R_i|^2]^{\frac{1}{2}},$$

where $|R_i|$ is the area of R_i, and then let

$$\Psi(f) = \sup \Psi(f;\rho) \text{ for all admissible subdivisions } \rho.$$

Since $\Psi(f) \leqq \Phi_1(f) + \Phi_2(f) + 1$, it follows that $\Psi(f) < +\infty$. For continuous surfaces, $\Psi(f)$ is the Geöcze area so that, in particular, for the integral means of our summable $f(x,y)$, we have $\Psi(f_n) = A(f_n)$.

Again using the fact that if $f(x)$ is of bounded generalized variation, then for every $\epsilon > 0$ there is a $\delta > 0$ such that if $S \subset E$ and $\nu(S) < \delta$, then $\phi(f;S) < \phi(f) - \epsilon$, where E is the set of points of approximate continuity of $f(x)$, we obtain an alternate definition for $\Psi(f)$. Indeed, as we shall prove, $\Psi(f) = \chi(f)$, where

$$\chi(f) = \sup \chi(f;\rho) \text{ for all admissible subdivisions } \rho,$$

and

$$\chi(f;\rho) = \sum_{i=1}^{m} [\{\chi_1(f;R_i)\}^2 + \{\chi_2(f;R_i)\}^2 + |R_i|^2]^{\frac{1}{2}},$$

where

$$\chi_1(f;R_i) = \int_{c_i}^{d_i} |f(b_i,y) - f(a_i,y)|\, dy$$

$$\text{and } \chi_2(f;R_i) = \int_{a_i}^{b_i} |f(x,d_i) - f(x,c_i)|\, dx,$$

the rectangle R_i being given by $a_i \leqq x \leqq b_i$, $c_i \leqq y \leqq d_i$.

Analogously, if

$$\chi_1(f;\rho) = \sum_{i=1}^{m} \chi_1(f;R_i), \qquad \chi_2(f;\rho) = \sum_{i=1}^{m} \chi_2(f;R_i),$$

let

$$\chi_1(f) = \sup \chi_1(f;\rho), \qquad \chi_2(f) = \sup \chi_2(f;\rho)$$

for all admissible subdivisions. Obviously, we have

$$\chi_i(f) \leqq \Phi_i(f), \quad i = 1,2, \qquad \chi(f) \leqq \Psi(f).$$

We show that $\Psi(f) = \lim\limits_{n\to\infty} \Psi(f_n) = \lim\limits_{n\to\infty} A(f_n)$. We first let $\epsilon > 0$ and let ρ be an admissible subdivision $R_1, R_2, \cdots, R_m$ for which $\Psi(f;\rho) > \Psi(f) - \epsilon$. By Theorem 2,

$$\liminf_{n\to\infty} \Phi_1(f_n;R_i) \geqq \Phi_1(f;R_i) \text{ and } \liminf_{n\to\infty} \Phi_2(f_n;R_i) \geqq \Phi_2(f;R_i),$$

for every $i = 1, 2, \cdots, m$, so that $\liminf\limits_{n\to\infty} \Psi(f_n;R_i) \geqq \Psi(f;R_i)$. It follows that $\liminf\limits_{n\to\infty} \Psi(f_n;\rho) \geqq \Psi(f;\rho) \geqq \Psi(f) - \epsilon$ and, since $\epsilon > 0$ is arbitrary. $\liminf\limits_{n\to\infty} A(f_n) \geqq \Psi(f)$.

Conversely, $\limsup\limits_{n\to\infty} A(f_n) \leqq \chi(f)$. In showing this, we use the expression (3) for $A(f_n)$. Then

$$\begin{aligned}A(f_n)=\int_0^{1-1/n}\int_0^{1-1/n}&[1+\{n^2\int_0^{1/n}(f(x+1/n,y+v)-f(x,y+v))dv\}^2\\&+\{n^2\int_0^{1/n}(f(x+u,y+1/n)-f(x+u,y))du\}^2]^{\frac{1}{2}}\,dx\,dy\\=n^2\int_0^{1/n}\int_0^{1/n}dx\,dy\sum_{j=1}^{n-1}\sum_{i=1}^{n-1}&[n^{-4}+\{\int_0^{1/n}(f(x+\frac{i}{n},y+\frac{(j-1)}{n}+v)\\&-f(x+\frac{(i-1)}{n},y+\frac{(j-1)}{n}+v))dv\}^2\\+\{\int_0^{1/n}(f(x+\frac{(i-1)}{n}+u,y+\frac{j}{n})&-f(x+\frac{(i-1)}{n}+u,y+\frac{(j-1)}{n}))du\}^2]^{\frac{1}{2}}.\end{aligned}$$

But, for almost all $(x,y)\in[0,1-1/n]\times[0,1-1/n]$, the subdivision obtained by means of the lines

$$x=x,x+1/n,\cdots,x+(n-1)/n;\qquad y=y,y+1/n,\cdots,y+(n-1)/n,$$

is admissible, so that

$$\begin{aligned}\sum_{j=1}^{n-1}\sum_{i=1}^{n-1}[n^{-4}+\{\int_0^{1/n}&(f(x+\frac{i}{n},y+\frac{(j-1)}{n}+v)\\&-f(x+\frac{(i-1)}{n},y+\frac{(j-1)}{n}+v))dv\}^2\\+\{\int_0^{1/n}&(f(x+\frac{(i-1)}{n}+u,y+\frac{j}{n})\\&-f(x+\frac{(i-1)}{n}+u,y+\frac{(j-1)}{n}))du\}^2]^{\frac{1}{2}}\leqq\chi(f).\end{aligned}$$

It then follows, since $A(f_n)$ is the mean value of the left hand side of this last inequality, that $A(f_n)\leqq\chi(f)$, for every $n=1,2,\cdots$ so that $\limsup_{n\to\infty}A(f_n)\leqq\chi(f)$. Thus we have

$$\limsup_{n\to\infty}A(f_n)\leqq\chi(f)\leqq\Psi(f)\leqq\liminf_{n\to\infty}A(f_n),$$

and hence

$$\lim_{n\to\infty}A(f_n)=\chi(f)=\Psi(f).$$

Analogously, we can prove that $\chi_1(f)=\Phi_1(f)$, $\chi_2(f)=\Phi_2(f)$. This completes the proof of

THEOREM 3. *Every summable function $f(x,y)$ has the property R.*

According to L. Cesari [2], the function $f(x,y)$ is said to be of type gBVT, or of generalized bounded variation in the sense of Tonelli, if when a certain set of measure zero is neglected, the function is of bounded variation with respect to each variable on the remaining set for almost all values of the other variable, and these modified variation functions are both summable. As proved in [3], the summable function $f(x,y)$ is gBVT if and only if $\chi_1(f)$, $\chi_2(f)$ are both finite. The considerations above prove also that $f(x,y)$ is gBVT if and only if $\Phi_1(f)$ and $\Phi_2(f)$ are both finite.

We turn now to parametric surfaces. C. B. Morrey [4] obtained results regarding the integral means for a parametric representation $f: x_i(u,v)$, $i=1,2,3$; perhaps his main result is that if the $x_i(u,v)$ are all ACT and if each pair of partial derivatives $\partial x_i/\partial u$, $\partial x_i/\partial v$ appearing as a product in one of the jacobians belongs to conjugate L_p spaces, then the jacobians of the integral means $x_i^n(u,v)$ of $x_i(u,v)$ converge almost everywhere, and in L_1 to those of $x_i(u,v)$. In particular, it follows that f has property R. The following fact is important in this connection: If $f(x,y)$ is ACT and if $f_n(x,y)$ is the n-th integral mean of $f(x,y)$, then $\partial f_n(x,y)/\partial x$ and $\partial f_n(x,y)/\partial y$ are the n-th integral means of $\partial f(x,y)/\partial x$ and $\partial f(x,y)/\partial y$ almost everywhere. We observe that this holds for $f(x,y)$ of type gACT, but the proof is not very different from the one for the continuous case, and so it will not be given. The definition of a gACT function [3] is the same as that of an ACT function except that the continuity restriction is waived.

The following extension of Morrey's theorem holds.

THEOREM A. *A mapping* $f: x_i(u,v)$, $i=1,2,3$, *has property* R *if the* $x_i(u,v)$ *are gACT and one of the following conditions is satisfied:*

a) *All partial derivatives* $\partial x_i/\partial u$, $\partial x_i/\partial v$, $i=1,2,3$, *are in* L_2.

b) *At least two of the* $x_i(u,v)$ *have bounded partial derivatives and the others are integrable.*

c) *Each pair* $\partial x_i/\partial u$, $\partial x_i/\partial v$ *appearing as a product in one of the jacobians belongs to conjugate* L_p *spaces.*

It is understood that sets of measure zero may be neglected in calculating partial derivatives appearing in this theorem. The proofs are then much like the continuous case, and are omitted.

In conclusion, we wish to point out that there are mappings which do not satisfy any of the above conditions and still have property R. For example, we consider the Geöcze mapping $f: x_i(u,v)$, $i=1,2,3$. Here, let $f_1(x)$, $f_2(x)$ be a continuous mapping on $[0,1]$ whose image covers the unit square; and let $x_1(u,v)=f_1(u)$, $x_2(u,v)=f_2(u)$, $x_3(u,v)=0$. The functions $x_1(u,v)$,

$x_2(u,v)$ are not of type gACT. We observe that the integral means $x_3{}^n(u,v)$ are zero so that

$$A(f_n)=\int_0^{1-1/n}\int_0^{1-1/n}|\,\partial(x_1{}^n,x_2{}^n)/\partial(u,v)\,|\,du\,dv.$$

But $x_i(u,v)=x_i(u+h,v)$, $i=1,2$, for every u,v,h, so that

$$\int_0^{1/n}\int_0^{1/n}x_i(u+\xi,v+\eta)\,d\xi\,d\eta=\int_0^{1/n}\int_0^{1/n}x_i(u+h+\xi,v+\eta)\,d\xi\,d\eta,$$

and $x_i{}^n(u+h,v)=x_i{}^n(u,v)$. It follows that $\partial x_i{}^n/\partial u=0$, $i=1,2$; $n=1,2,\cdots$; u and v arbitrary. Thus $A(f_n)=0$, so that f has property R.

THE UNIVERSITY OF OKLAHOMA.

BIBLIOGRAPHY.

[1] T. Radó, "Sur le calcul de l'aire des surfaces courbes," *Fundamenta Mathematicae*, vol. 10 (1927), pp. 197-210.

[2] L. Cesari, "Sulle funzioni a variazione limitata," *Annali Scuola Normale Superiore Pisa*, Ser. 2, vol. 5 (1936), pp. 299-313.

[3] C. Goffman, "Lower semi-continuity and area functionals, I. The non-parametric case," *Rendiconti del Circolo Matematico di Palermo*, Ser. 2, vol. 2 (1953), pp. 203-235.

[4] C. B. Morrey, "A class of representations of manifolds," *American Journal of Mathematics*, vol. 57 (1935), pp. 275-293.

[5] T. Radó, *Length and Area*, New York, 1948.

[6] L. Cesari, *Surface Area*, Princeton, in press.

ON THE RAMIFICATION OF ALGEBRAIC FUNCTIONS.*[1]

By SHREERAM ABHYANKAR.

Introduction. The investigations of the present paper arose from our (unsuccessful) attempt to adapt the ideas of Jung's classical proof of the theorem of local uniformization on algebraic surfaces over the field of complex numbers[2] to the solution of the local uniformization problem for algebraic surfaces over fields of nonzero characteristic.[3] To explain matters, let us consider the following situation. Let T be a rational transformation from an r-dimensional normal algebraic variety V onto an r-dimensional nonsingular algebraic variety W such that T and T^{-1} are free from fundamental points. Let the ground field k be algebraically closed of characteristic p. Let P and Q be corresponding points on V and W respectively and let B be the branch locus of T^{-1} on W. Let E and F be the quotient fields of the completions of the quotient rings of P and Q on V and W respectively. Then E can be canonically considered to be a finite separable algebraic extension of F. Let E^* be a least Galois extension of F containing E. Let us denote by $G(P/Q)$ the local Galois group of P over Q, i.e., the Galois group of E^* over F.

The basic facts underlying Jung's proof, then, are that for $p=0$ and $r=2$ we have: (1) If Q is a simple point of B then $G(P/Q)$ is cyclic and P is simple for V. (2) If Q is an ordinary double point of B then $G(P/Q)$ is a direct product of two cyclic groups. We shall give algebraic proofs of (1) and (2) in Section 2, and shall show in Section 3, by examples, that these statements are no longer true if $p \neq 0$; namely we shall show that P may be a singular point of V even if Q is a simple point of B, and that the local Galois groups above simple or ordinary double points of B can be very complicated and even insolvable. The collapse of statements (1) and (2) is the reason for the nonadaptability of Jung's method of uniformization to the case

---

* Received April 6, 1955.

[1] The material of this paper together with paper [1] forms the author's Ph. D. thesis at Harvard University. The work was supported by a research project at Harvard University sponsored by the Office of Ordnance Research, U. S. Army, under Contract DA-19-020-ORD-3100.

[2] See [4]. For a simplified version of Jung's proof see [10].

[3] We have proved the theorem of local uniformization on algebraic surfaces over fields of nonzero characteristic in [1].

of ground fields of nonzero characteristic. An algebraic treatment of the main parts of Jung's proof will be given in Section 4.

Let now $p \neq 0$ and assume that B has a t-fold normal crossing at Q with $t \leqq r$ (for definition see Section 3). Let $\pi(P/Q)$ be the smallest subgroup of $G(P/Q)$ containing all the p-Sylow subgroups of $G(P/G)$. Then, as will be proved in Section 2, the factor group $G(P/Q)/\pi(P/Q)$ is a direct product of t cyclic groups. This characterization of local Galois groups above normal crosings is the main result of this paper. Also we shall include in Section 1, Zariski's own proof of his theorem about the purity of the branch locus.

My warmest thanks are due to Professor O. Zariski for his kind encouragement and valuable advice.

Notations. We shall use the following notations. For a valuation v of a field L we shall denote by R_v the valuation ring of v and by M_v the maximal ideal in R_v. We shall say that an integral domain A is normal, if A is integrally closed in its quotient field. If R is a local ring and M its maximal ideal, we shall express this by saying, "(R, M) is a local ring." If R is the quotient ring of an irreducible subvariety of an algebraic variety, we shall say that R is algebraic. If W is an irreducible subvariety of an algebraic variety V, we shall denote by $Q(W, V)$ the quotient ring of W on V and by $M(W, V)$ the maximal ideal in $Q(W, V)$. When the reference to the embedding variety V is clear we shall simply write $Q(W)$ and $M(W)$ for $Q(W, V)$ and $M(W, V)$ respectively. Unless otherwise stated, by a point we shall mean a rational point over the ground field under consideration.

1. The branch locus.

The basic definitions and some preliminary results needed in the present paper were given by us in Section 2 of [1]. In this section we add a few more preliminary results.

In the following two lemmas we collect some well known facts about derived normal models. For proof of these lemmas we refer to Zariski's papers: see the remarks on pages 68-70 of [13] and the discussion of the birational case on pages 506-511 of [11].

LEMMA 1. *Let K be an r-dimensional algebraic function field over an algebraically closed ground field k, and let K^* be a finite separable extension of K. Let V and V^* be normal projective models respectively of K/k and K^*/k such that the rational transformation T from V^* onto V (determined by the embedding $K \subset K^*$ of the function fields of V and V^*) and its inverse map T^{-1} are free from fundamental points. Then:* (a) *To an irreducible* *subvariety* W^* *of* V^* *corresponds a unique irreducible subvariety* W *of* V. *Furthermore* $\dim W^* = \dim W$ *and* $Q(W^*) \supset Q(W)$. (b) *To an irreducible subvariety* W *of* V *correspond a finite number of irreducible subvarieties* $W^*_1, W^*_2, \cdots, W^*_s$ *of* V^*. *Let* (R, M) *and* (R^*_i, M^*_i) *be the quotient rings of* W^*_i *and* W *respectively. Then* $R^*_1, R^*_2, \cdots, R^*_s$ *are exactly the local rings in* K^* *lying above* R.

LEMMA 2. *Given* K/k, K^* *and* V *as in Lemma* 1, *there exists* V^* *as in Lemma* 1. *Furthermore,* V^* *is unique up to a biregular transformation.*

DEFINITION 1. *We call* V^* *a derived normal model of* V *in* K^*/k. *Now let the notation be as in part* (b) *of Lemma* 1. *For a fixed index* a, *let* E *and* E^*_a *be the quotient fields of the completions of* R *and* R^*_a *respectively; and embed* E *canonically in* E^*_a. *Recall from Section* 2 *of* [1] *that:*

$$d(M^*_a : M) = [E^*_a : E], \quad g(M^*_a : M) = [(R^*_a/M^*_a) : (R/M)]_s,$$

$$i(M^*_a : M) = [(R^*_a/M^*_a) : (R/M)]_i, \quad r(M^*_a : M) = d(M^*_a : M) g(M^*_a : M)^{-1},$$

$$\bar{r}(M^*_a : M) = r(M^*_a : M) i(M^*_a : M)^{-1}.$$

Now we define:

$d(W^*_a : W) =$ *degree of* W^*_a *over* $W = d(M^*_a : M)$;

$g(W^*_a : W) =$ *separable residue degree of* W^*_a *over* $W = g(M^*_a : M)$;

$i(W^*_a : W) =$ *inseparability index of* W^*_a *over* $W = i(M^*_a : M)$;

$r(W^*_a : W) =$ *ramification index of* W^*_a *over* $W = r(M^*_a : M)$;

$\bar{r}(W^*_a : W) =$ *reduced ramification index of* W^*_a *over* $W = \bar{r}(M^*_a : M)$.

We shall say that W^*_a *is ramified over* W*—or that* W^*_a *is ramified for the transformation* T *(or* T^{-1}*)—if* M^*_a *is ramified over* M. *Also we shall say that* W *is a branch variety of* V *for the transformation* T *(or* T^{-1}*) if* M *is a branch ideal in* R *for the extension* $K \subset K^*$.

Then W is a branch variety of V for the transformation T (or T^{-1}) if $r(W^*_i : W) > 1$ for some i or equivalently if $\sum_{i=1}^{s} g(W^*_i : W) < [K^* : K]$. By the *branch locus* of T or T^{-1}, denoted by $B(T)$ or $B(T^{-1})$ we shall mean the set of all (rational) points P of V/K such that P is on some irreducible subvariety of V/k which is a branch subvariety for T.

LEMMA 3. *Let* K/k, K^*, V, V^* *and* T *be as in Lemma* 1. *Then an* *irreducible subvariety W of V is a branch subvariety if and only if $W \subset B(T)$. Furthermore, $B(T)$ is a (proper) subvariety of V.*

Proof. Let A be an affine coordinate ring of V and let A^* be the integral closure of A in K^*. Let us denote by $D(A^*/A)$ the ideal in A generated by all the discriminants $d(w_1, w_2, \cdots, w_n)$ over K of K-bases $(w_1, w_2, \cdots, w_n)$ of K^* which belong to A^*. Let $U \subset W$ be two irreducible proper subvarieties of V which are at finite distance for A. Then W is a branch variety of T if and only if $D(A^*/A) \subset M(W) \cap A$; see [7]. Therefore, if W is a branch variety then $D(A^*/A) \subset M(W) \cap A \subset M(U) \cap A$ and hence U is a branch variety. Again, if W is not a branch variety then $D(A^*/A) \not\subset M(W) \cap A$ and hence for some maximal ideal N in A containing $M(W) \cap A$ we must have: $D(A^*/A) \not\subset N$ (Hilbert Nullstellensatz). This proves our first assertion. To prove the second assertion let A be any affine coordinate ring of V and let A^* be the integral closure of A in K^*. Since K^*/K is separable, $D(A^*/A) \neq 0$ and hence the set of all points of $B(T)$ which are at finite distance for A, is a proper subvariety of the affine part of V given by A; i.e. "$B(T)$ is a set on a pure r-dimensional variety V/k such that if H is any hyperplane defined over k in the embedding projective space S_m/k of V then $B(T) \cap (S_m - H)$ is a proper subvariety in the affine space $S_m - H$." By induction on r, one can easily show that the assumption between quotation marks implies that $B(T)$ is a proper subvariety of V. Our second assertion can be proved also by using the concept and simple properties of symmetric products developed by Zariski in [14]. The proof, then, can be given as follows: Let P be a general point of V/k and let $P_1, P_2, \cdots, P_n$ be the n points which correspond to P for T^{-1} (taking coordinates in some fixed universal domain. Let $\bar{P} = P_1 \circ P_2 \circ \cdots \circ P_n$ be the symmetric product of $P_1, P_2, \cdots, P_n$; and let $\bar{T}$ be the irreducible algebraic correspondence with a general point pair $(P, \bar{P})$ over k. Since T is a finitely valued rational transformation without fundamental points, the same is true for $\bar{T}$; ($\bar{T}$ is in fact birational). Hence for any point Q (with coordinates in a universal domain) of V there corresponds a unique point $\bar{Q} = Q_1 \circ Q_2 \circ \cdots \circ Q_n$ under $\bar{T}$ and hence the singular locus S of $\bar{T}$ is exactly the set of points on V to which correspond less than n points on V^* for T^{-1}. Therefore $B(T)$ coincides with the set of rational points of S (since k is algebraically closed), and hence $B(T)$ is a proper subvariety of V.

DEFINITION 2. *Let v be a real discrete valuation of a field K, K^* a finite separable extension of K, and v^* a K^*-extension of v. Let $M = M_v$ and $M^* = M_{v^*}$. We define:*

$d(v^*:v)$ = *degree of* v^* *over* $v = d(M^*:M)$;

$i(v^*:v)$ = *inseparability index of* v^* *over* $v = i(M^*:M)$;

$g(v^*:v)$ = *separable residue degree of* v^* *over* $v = g(M^*:M)$;

$r(v^*:v)$ = *ramification index of* v^* *over* $v = r(M^*:M)$;

$\bar{r}(v^*:v)$ = *reduced ramification index of* v^* *over* $v = \bar{r}(M^*:M)$.

Remark. In the notation of Lemma 3, let U be an irreducible $r-1$ dimensional subvariety of V and $U^*_1, U^*_2, \cdots, U^*_s$ the corresponding subvarieties of V^*. Let u and $u^*_1, u^*_2, \cdots, u^*_s$ be the real discrete valuations of K and K^* respectively such that $Q(U) = R_u$ and $Q(U^*_j) = R_{u^*_j}$. We observe that $\bar{r}(u^*_j:u)$ = reduced ramification index (in our sense) of u^*_j over u = the usual ramification index (in the sense of valuation theory) of u^*_j over u. One might be tempted to define: U is a branch variety if and only if $\bar{r}(u^*_j:u) > 1$ for some j. One can easily see that with this definition the above Lemma is false, for then U may consist entirely of branch points without itself being a branch variety—this will certainly be so if $\bar{r}(u^*_j:u) = 1$ for all j and $i(u^*_j:u) > 1$ for some j; see Example 1 of Section 3 (the same example also shows that with this definition Theorem 1, which we shall prove in a moment, is false). Therefore our definition of a branch variety is the right one. Of course for characteristic zero the two definitions coincide.

LEMMA 4. *Let* K/k *and* V *be as in Lemma* 1. *Let* W *a proper irreducible subvariety of* V, *and* P *a simple point of* V *which lies on* W. *Let* (o, m) *and* (R, M) *be the respective local rings of* W *and* P *on* V. *Let* $x_1, x_2, \cdots, x_r$ *be regular parameters in* R. *Then* $x_1, x_2, \cdots, x_r$ *are separating coordinates for* W, *i.e., the set* $(x_1, x_2, \cdots, x_r)$ *satisfies the following four conditions:* (1) $(x_1, x_2, \cdots, x_r)$ *is a separating transcendence basis of* K/k, (2) $k[x_1, x_2, \cdots, x_r] \subset o$, (3) $k[x_1, x_2, \cdots, x_r]$ *contains a basis of* m, *and* (4) o/m *is a finite separable algebraic extension of*

$$k[x_1, x_2, \cdots, x_r]/(k[x_1, x_2, \cdots, x_r] \cap m).$$

Proof. Let $(\bar{R}, \bar{M})$ be the completion of (R, M). Since $\bar{R}$ is regular and since it contains a representative k of its residue field, it follows as in Section 5 of [6] that we can indentify $\bar{R}$ with $k[[x_1, x_2, \cdots, x_r]]$, and we can consider K as a subfield of $k((x_1, x_2, \cdots, x_r))$. By Prop. 1.5 of [4], $k((x_1, x_2, \cdots, x_r))$ and $k(x_1, x_2, \cdots, x_r)^{1/p}$ are linearly disjoint over $k(x_1, x_2, \cdots, x_r)$, and hence $K/k(x_1, x_2, \cdots, x_r)$ is separable.[4]

---

[4] p denotes the characteristic of k. We are assuming $p \neq 0$; for in case $p = 0$, $K/k(x_1, x_2, \cdots, x_r)$ is trivially separable.

Let $S = k[x_1, x_2, \cdots, x_r]$, $N = (x_1, x_2, \cdots, x_r)S$, $R_1 = S_N$, and $M_1 = NR_1$. Let $\bar{R}_1$ be the integral closure of R_1 in K. Then as in Proposition 2 of [1], $R = (\bar{R}_1)_{\bar{R}_1 \cap M}$. Let $A = S \cap m$, $o_1 = S_A$, and $m_1 = Ao_1$. Let $\bar{o}_1$ be the integral closure of o_1 in K. Then $o = (\bar{o}_1)_{\bar{o}_1 \cap m}$. By part (III), Lemma 4 of [1], $r(M : M_1) = 1$ Therefore by part (a) of Lemma 6 of [1], $r(m : m_1) = 1$. This shows that $(x_1, x_2, \cdots, x_r)$ is a separating coordinate system of W.

The following theorem and its proof are due to Professor Zariski and were communicated to me personally. The proof is based on Zariski's paper [12].

THEOREM 1. *Let K/k, K^*, V, V^* and T be as in Lemma 1. Let $B = B(T)$ be the branch locus of T on V and let $B^* = T^{-1}\{B\}$. Let W be a simple irreducible subvariety of V and W^* a corresponding subvariety of V^*.* (a) *If W^* is ramified over W, then there exists an $r - 1$ dimensional irreducible subvariety U^* of V^* containing W^* such that U^* is ramified over its corresponding subvariety U of V.* (b) *The components of B^* and B passing respectively through W^* and W are pure $r - 1$ dimensional.* (c) *If $B(T)$ contains no singular points of V, in particular if V is nonsingular, then $B(T)$ is pure $r - 1$ dimensional.*

Proof. First observe that (b) and (c) follow at once from (a) and that in the proof of (a) we may, without loss of generality, let W^* and W be points (rational over k) P^* and P respectively. Now let (R^*, M^*) and (R, M) be the respective local rings of P^* and P, and let $(x_1, x_2, \cdots, x_r)$ be regular parameters in R. Let $(y_1, y_2, \cdots, y_m)$ be an affine general point of V^*/k for which P^* is at finite distance. Let $A = k[y_1, y_2, \cdots, y_{m+r}]$ where $y_{m+i} = x_i$ for $i = 1, 2, \cdots, r$. Then $R^* = A_{A \cap M^*}$ and hence after replacing V^* by the variety having $(y_1, y_2, \cdots, y_{m+r})$ as a general point over k we may assume that $(y_1, y_2, \cdots, y_{m+r})$ is a general point of V^*/k. Let $L = k(x_1, x_2, \cdots, x_r)$ and let v^* be the zero dimensional irreducible variety in the projective space S_m/L with general point $(y_1, y_2, \cdots, y_m)$ over L. Then $Q(v^*, S_m/L) = Q(V^*, S_{m+r}/k)$.[5] Now v^* is a simple point of S_m/L and hence $M(V^*, S_m/L)$ has a basis $(f_1, f_2, \cdots, f_m)$ of m elements. By multiplying by a suitable unit in $Q(V^*, S_m/L)$ we may assume that $f_i = F_i(y_1, y_2, \cdots, y_{m+r})$ where $F_i(Y_1, Y_2, \cdots, Y_{m+r}) \varepsilon k[Y_1, Y_2, \cdots, Y_{m+r}]$ for $i = 1, 2, \cdots, m$. By Lemma 4, K^*/L is separable and hence $J \neq 0$ at $Y_i = y_i$ $(i = 1, 2, \cdots, m + r)$ where $J = \partial(F_1, F_2, \cdots, F_m)/\partial(Y_1, Y_2, \cdots, Y_m)$.[6]

---

[5] See Section 2.2 of [12].

[6] See the Criterion for Uniformizing Parameters on page 27 of [12].

Suppose, if possible, that $J \neq 0$ at P^*. Then

$$\partial(F_1, F_2, \cdots, F_{m+r})/\partial(Y_1, Y_2, \cdots, Y_{m+r}) = J \neq 0$$

at P^*, where $F_{m+i} = Y_{m+i}$ for $i = 1, 2, \cdots, r$. Therefore $F_1, F_2, \cdots, F_{m+r}$ are uniformizing parameters in $Q(P^*, S_{m+r}/k)$.[6] Let

$$E = k[Y_1, Y_2, \cdots, Y_{m-r}] \cap M(V^*, S_{m+r}/k).$$

Since $R^* = Q(P^*, S_{m+r}/k)/E$ and since $F_i \varepsilon E$ for $i = m+1, m+2, \cdots, m+r$, we conclude that $M^* = (x_1, x_2, \cdots, x_r)R^*$ i.e. P^* is unramified over P, which is a contradiction. Therefore $J = 0$ at P^*, i.e. $J^* \varepsilon M^*$ where J^* is the value of J at $Y_1 = y_1$, $Y_2 = y_2, \cdots, y_{m+r} = y_{m+r}$.

Let N^* be an isolated prime ideal of J^*R^* in R^*. Since J^*R^* is principal, N^* is a minimal prime ideal in R^* and hence defines an $r-1$ dimensional subvariety U^* of V^* through P^*.[7] Let $N = N^* \cap R$ and U the subvariety defined by N. Let $S = k[x_1, x_2, \cdots, x_r]$, $D = S \cap N$. Since k is algebraically closed, S_D/DS_D is separably generated over k. By Maclane's theorem (8.2 of [12]) we can choose a separating basis from the residues of $x_1, x_2, \cdots, x_r$; say for instance $x_2, x_3, \cdots, x_r$ do this job. Fix

$$G(Y_{m+1}, Y_{m+2}, \cdots, Y_{m+r}) \varepsilon k[Y_{m+1}, Y_{m+2}, \cdots, Y_{m+r}]$$

such that $D = G(x_1, x_2, \cdots, x_r)S$. As in the proof of Lemma 4, $N = G(x_1, x_2, \cdots, x_r)R$ and $R_N/(NR_N)$ is a separable extension of $k(x_2, x_3, \cdots, x_r)$. Suppose, if possible, that N^* is unramified over N. Then $N^* = G(x_1, x_2, \cdots, x_r)R^*$ and $R^*_{N^*}/(N^*R^*_{N^*})$ is a separable extension of $k(x_2, x_3, \cdots, x_r)$. Therefore $F_1(Y_1, Y_2, \cdots, Y_{m+r})$, $F_2(Y_1, Y_2, \cdots, Y_{m+r})$, $\cdots, F_m(Y_1, Y_2, \cdots, Y_{m+r})$, $G(Y_{m+1}, Y_{m+2}, \cdots, Y_{m+r})$ are regular parameters in $Q(U^*, S_{m+r}/k)$; (see corollary on page 9 of [12]). Therefore considering G as a point in $S_{m+1}/k(x_1, x_2, \cdots, x_{r-1})$, we have:

$$\partial(F_1, F_2, \cdots, F_m, G)/\partial(Y_1, Y_2, \cdots, Y_{m+1}) = J\frac{\partial G}{\partial Y_{m+1}} \neq 0$$

at $Y_1 = y_1$, $Y_2 = y_2, \cdots, Y_{m+r} = y_{m+r}$.[6] Therefore $J \neq 0$ at $Y_1 = y_1$, $Y_2 = y_2$, $\cdots, Y_{m+r} = y_{m+r}$, which is a contradiction. Therefore U^* is ramified over U.[8]

PROPOSITION 1. *Let (R, M) be a normal algebraic local ring with quotient field K. Let K' be a finite separable extension of K and let K^* be a least Galois extension of K containing K'. Then M is a branch ideal (in R) for the extension K'/K if and only if it is so for the extension K^*/K.*

---

[7] R^* being normal all the primes of J^*R^* are minimal.

[8] $J = 0$ is a local equation of the component of B^* through P^*.

Proof. If M is not a branch ideal for K^*/K then by Lemma 4 of [1] it follows that M is not a branch ideal for K'/K. Now assume that M is not a branch ideal for K'/K. Let (R^*_j, M^*_j) for $j=1,2,\cdots,t$ be the local rings in K^* lying above R. Let G be the Galois group of K^*/K, G' the Galois group of K^*/K', and G_j the inertia group of M_j over M. Then by Lemma 5.V.1 of [1], $G' \supset G_j$ for $j=1,2,\cdots,t$. Let $\bar{G}$ be the smallest subgroup of G which contains $G_1, G_2, \cdots, G_t$. Then $\bar{G}$ is a normal subgroup of G and $\bar{G} \subset G'$. Therefore the fixed field $\bar{K}$ of $\bar{G}$ is a Galois extension of K containing K', and hence $\bar{K}=K^*$, i.e., $\bar{G}=(1)$. Therefore $G_1=G_2=\cdots=G_t=(1)$, and hence by Lemma 5.I.3 of [1], $r(M^*_j:M)=1$ for $j=1,2,\cdots,t$; i.e., M is not a branch ideal for K^*/K.

2. Local Galois groups.

Definition 3. *Let G be a finite group, p a prime number and s an integer. Let π be the smallest subgroup of G which contains all the Sylow p-groups of G. If G is an abelian group which can be expressed as a direct product of t cyclic subgroups with $t \leqq s$, then we shall say that G is a 0_s-group. If G/π is a 0_s-group, we shall say that G is a p_s-group.*

Definition 4. *Let (R,M) be the local ring of a simple point P of an r dimensional algebraic variety V, and let W be a pure $r-1$ dimensional subvariety of V. Let $W_1, W_2, \cdots, W_s$ be the irreducible components of W which pass through P. Let $N_i = R \cap M(W_i, V)$. Let s be an integer with $s \leqq r$. Then we shall say that W has an s-fold normal crossing at P if there exists a minimal basis $(x_1, x_2, \cdots, x_r)$ of M such that $x_iR=N_i$ for $i=1,2,\cdots,s$.*

Lemma 5. *Let $R=k[[X_1,X_2,\cdots,X_r]]$, $E=k((X_1,X_2,\cdots,X_r))$, and E^* the root field of $f(Z)=Z^n-X$ over E, where in case of nonzero characteristic we assume that n is not divisible by the characteristic. Let R^* be the integral closure of R in E^*, and let X^* be a root of $f(Z)$ in E^*. Then $R^*=\sum_{i=0}^{n-1} X^{*i}R$, R^* is a regular local ring of dimension r and $(X^*, X_2, X_3, \cdots, X_r)$ are regular parameters in R^*.*

Proof. Let v be the valuation of E such that $R_v=R_{(X_1R)}$, and let v^* be a E^*-extension of v. Since $nv^*(X^*)=v(X_1)$, v^* is the only extension of v to E^*. Since X_1^{n-1} is the discriminant of $f(Z)$, we have that: $R^*=\sum_{i=0}^{n-1}(X^{*i}X_1^{1-n})R$, (see page 79 of [9]). $u^* \varepsilon R^*$ implies that $u^*=\sum_{i=0}^{n-1} X^{*i}X_1^{1-n}u_i$ with $u_i \varepsilon R$, and $v^*(u^*) \geqq 0$. The v^* values of any two different nonzero terms in the above sum are distinct. Hence $u_i \neq 0$ implies that $v^*(X^{*i}X_1^{1-n}u_i) \geqq 0$ and hence that $X_1^{1-n}u_i \varepsilon R$. Therefore $R^* = \sum_{i=0}^{n} X^{*i}R$. The remaining part of the lemma follows at once from this.

In the rest of this section we shall deal with the following situation. Let K be an r dimensional function field over an algebraically closed ground field k of characteristic p and let K^* be a finite separable extension of K. Let V be a projective normal model of K/k, V^* a derived normal model of V in K^*, and T the rational transformation from V^* onto V. Let P and P^* be corresponding points of V and V^*, and let (R,M) and (R^*,M^*) be the local rings of P and P^* respectively. Let Ω be an algebraic closure of K^*.

LEMMA 6. *Assume K^*/K is Galois and let $y_1, y_2, \cdots, y_h$ be a finite number of elements in R such that $N_j = y_jR$ is a minimal prime ideal in R and hence that R_{N_j} is the valuation ring a real discrete valuation w_j of K. Let n_j be the reduced ramification index over w_j of a K^*-extension of w_j.*[9] *Assume that if $p \neq 0$ then $n_j \not\equiv 0(p)$ for $j = 1, 2, \cdots, h$. Let y_j^{1/n_j} be an n_j-th root of y_j in Ω. Let*

$$\bar{K} = K(y_1^{1/n_1}, y_2^{1/n^2}, \cdots, y_h^{1/n_h}) \text{ and } \bar{K}^* = K^*(y_1^{1/n_1}, y_2^{1/n_2}, \cdots, y_h^{1/n_h}).$$

Let N_0 be a minimal prime of R different from $N_1, N_2, \cdots, N_h$ and let w_0 be the valuation of K with valuation ring R_{N_0}. Let $\bar{w}^*_j$ be an extension of w_j to $\bar{K}^*$, w^*_j the K^*-restriction of $\bar{w}^*_j$, and $\bar{w}_j$ the $\bar{K}$-restriction of $\bar{w}^*_j$. Then: (a) *$\bar{K}^*/K$, $\bar{K}/K$, $\bar{K}^*/K^*$ and $\bar{K}^*/\bar{K}$ are Galois extensions.* (b) *$r(\bar{w}^*_j : w^*_j) = \bar{r}(\bar{w}^*_j : w^*_j) = 1$ and $r(\bar{w}_j = w_j) = \bar{r}(\bar{w}_j : w_j) = n_j$ for $j = 1, 2, \cdots, h$.* (c) *$r(\bar{w}^*_0 : w^*_0) = r(\bar{w}_0 : w_0) = 1$, $r(\bar{w}^*_0 : \bar{w}_0) = r(w^*_0 : w_0)$,* and *$\bar{r}(\bar{w}^*_0 : \bar{w}_0) = \bar{r}(w^*_0 : w_0)$.*

Proof. We shall assume that $h = 1$, since the general case follows from this by a straightforward induction. (a) follows from the fact that the compositum of two Galois extensions of K is again a Galois extension of K. Now let $[\bar{K}^* : K^*] = n$. Then $n_1 \equiv 0(n)$. Let $x = y_1^{1/n_1}$ and $z = x^n$. Then $f(X) = X^n - z$ is the minimal monic polynomial of x over K^*. Now z^{n-1} is the discriminant of $f(X)$. Since z^{n-1} is a unit in $R_{w^*_0}$, $r(\bar{w}^*_0 : w^*_0) = 1$. Fix $u \varepsilon R_{w^*_1}$ such that $uR_{w^*_1} = M_{w^*_1}$. Then x/u is a primitive element of $\bar{K}^*/K^*$, and $g(X) = X^n - zu^{-n}$ is the minimal monic polynomial of x/u over K^*. Since $w^*_1(zu^{-n}) = 0$, the w^*_1-residue e of zu^{-n} is a nonzero element. Hence the discriminant of $X^n - e$ is also a nonzero element of $R_{w^*_1}/M_{w^*_1}$.

---

[9] Since K^*/K is Galois, n_j depends only on w_j.

Therefore the discriminant of $g(X)$ is a unit in $R_{w^*_1}$, and hence $r(\bar{w}^*_1 : w^*_1) = 1$. Since $n_1\bar{w}_1(x) = w_1(y_1)$, we have that $r(\bar{w}_1 : w_1) = \bar{r}(\bar{w}_1 : w_1) = n_1$. The rest of the assertions follow in a similar fashion.

LEMMA 7. *Assume that P is simple; P^* is the only point which corresponds to P; $Z^n - x_1$ is the minimal equation of a primitive element x^* of K^*/K; $[K^* : K] \not\equiv 0$ if $p \neq 0$, and that x_1 is part of a minimal basis $(x_1, x_2, \cdots, x_r)$ of M. Then P^* is simple and $M^* = (x^*, x_2, x_3, \cdots, x_r)R^*$.*

Proof. Follows from Proposition 1 of [1] and Lemma 5.

THEOREM 2. *Let $\bar{B}$ be the branch locus on V. Assume that P is simple and that the component B of $\bar{B}$, which is ramified for the extension $P \rightarrow P^*$, has a t-fold normal crossing at P.*[10] *Then the local Galois group $G(P^*/P)$ is a p_t-group.*

Proof. Let K' be the least Galois extension of K such that $K^* \subset K' \subset \Omega$, P' a point corresponding to P^* on a derived normal model of V^* in K', and (R', M') the local ring of P'. Let K_s and K^*_s be the splitting fields of P' over P and P^* respectively, and let $\bar{K}'$ be the least Galois extension of K_s such that $K^*_s \subset \bar{K}' \subset K'$. Let P_s, P^*_s, $\bar{P}'$ be the points corresponding to P' on derived normal models V_s, V^*_s, $\bar{V}'$ of V in K_s, K^*_s, $\bar{K}'$ respectively. The results of Section 1 and of Section 2 of [1] imply the following: P_s is simple; if B_s is the image of B for the transformation $V \rightarrow V_s$, then B_s is the branch variety for the extension $P \rightarrow \bar{P}'$, B_s has a t-fold normal crossing at P_s; and $G(P'/P_s) = G(P^*/P)$. Therefore we may replace V, P, V^*, P^* by V_s, P_s $\bar{V}'$, $\bar{P}'$ respectively.

Hence to start with we may assume that P^* is the only point corresponding to P and that K^*/K is Galois. Let G be the Galois group of K^*/K and let π be the smallest subgroup of G which contains all the p-Sylow groups of G if $p \neq 0$ and let $\pi = (1)$ if $p = 0$. Let $\tilde{K}$ be the fixed field of π, $\tilde{V}$ a derived normal model of V in $\tilde{K}$ and P the point in $\tilde{V}$ corresponding to P^*. Then the branch locus $\tilde{B}$ for the extension $P \rightarrow \hat{P}$ is a subvariety of B and hence $\tilde{B}$ has a m-fold normal crossing at P with $m \leqq t$. We have to show that G/π is a 0_t-group, and hence enough to show that G/π is a 0_m-group.

Therefore to start with we may assume that if $p \neq 0$ then $[K^* : K] \not\equiv 0(p)$; and then we have to show that G is a 0_t-group. Let $(x_1, x_2, \cdots, x_r)$ be regular parameters in R such that $R_{(x_jR)} = Q(B_j, V)$ for $j = 1, 2, \cdots, t$ where

---

[10] By Theorem 1, $\bar{B}$ is pure $r - 1$ dimensional. B consists of those irreducible components B_j of $\bar{B}$ passing through P for which there exists a corresponding variety B^*_j through P^* with $r(B^*_j : B_j) > 1$.

$B_1, B_2, \cdots, B_t$ are the irreducible components of B. Let B^*_j be an irreducible subvariety of V^* corresponding to B_j, and let $n_j = r(B^*_j : B_j)$.[11] Let x_j^{1/n_j} be an n_j-th root of x_j in Ω. Let $K_2 = K^*(x_1^{1/n_1}, x_2^{1/n_2}, \cdots, x_t^{1/n_t})$ and $K_1 = K(x_1^{1/n_1}, x_2^{1/n_2}, \cdots, x_t^{1/n_t})$. Let P_2 be a point corresponding to P^* on a derived normal model V_2 of V in K_2 and let P_1 be the point corresponding to P_2 on a derived normal model V_1 of V in K_1. Let (R_1, M_1) be the quotient ring of P_1. Theorem 1 and Lemmas 6 and 7 imply that P_1 is a simple point, $M_1 = (x_1^{1/n_1}, x_2^{1/n_2}, \cdots, x_t^{1/n_t},\ x_{t+1}, x_{t+2}, \cdots, x_r)$, $G(P_2 : P) = G(P_1 : P) =$ the direct product of t cyclic groups of orders $n_1, n_2, \cdots, n_t$. Since $G(P^* : P)$ is a subgroup of $G(P_2 : P)$, we conclude that $G(P^* : P)$ is 0_t-group.

THEOREM 3.[12] *Let the assumptions be as in Theorem 2 and let* $p = 0$. *Let* $(x_1, x_2, \cdots, x_r)$ *be regular parameters in* R *such that* $R_{(x_jR)} = Q(B_j, V)$ *for* $j = 1, 2, \cdots, t$; *where* $B_1, B_2, \cdots, B_t$ *are the irreducible components of* B. *Let* $\bar{R} = k[[x_1, x_2, \cdots, x_r]]$ *and* $\bar{R}^*$ *be the completions of* R *and* R^* *respectively; and embed* $\bar{R}$ *in* $\bar{R}^*$ *canonically. Let* E *and* E^* *be the respective quotient fields of* $\bar{R}$ *and* $\bar{R}^*$. *Then* P^* *is Galois over* P, $G(P^* : P)$ *is a* 0_t*-group and* $E^* \subset k((x_1^{1/n_1}, x_2^{1/n_2}, \cdots, x_t^{1/n_t}, x_{t+1}, x_{t+2}, \cdots, x_r))$ *where* $n_1, n_2, \cdots, n_t$ *are suitable integers.*

Proof. Follows from the considerations made in the proof of Theorem 2.

THEOREM 4. *Let the assumptions be as in Theorem 3 and let* $t = 1$, *i.e., let* B *have a simple point at* P. *Then* P^* *is simple,*

$$\bar{R}^* = k[[x_1^{1/n_1}, x_2, \cdots, x_r]],$$

and $x_1 R^*$ *is primary, i.e.,* B *does not split locally for the extension* $P \rightarrow P^*$.

Proof. That P^* is simple and $R^* = k[[x_1^{1/n_1}, x_2, x_3, \cdots, x_r]]$ follows from Theorem 3. Since x_1 is the power of an irreducible element in $\bar{R}^*$, it must already be so in R^*.

Remark. Let $G(P^* : P)$ be the local Galois group over a t-fold normal crossing as referred to in Theorem 2. If $p \neq 0$, $G(P^* : P)$ can be very messy and insolvable even for $t = 2$ and $t = 1$ as will be shown by Examples 3 and 4 in Section 3. Hence in a certain sense, Theorem 2 gives a best possible description of local Galois groups over normal crossings. Since

---

[11] $r = \bar{r}$ since $[K^* : K] \not\equiv 0(p)$ if $p \neq 0$.

[12] Known in the classical case when k is the field of complex numbers; the known proofs being topological.

Theorem 4 essentially depends on the fact that for $p=0$ the local Galois group above a simple point of the branch locus is cyclic and since for $p\neq 0$ such a local Galois group need not be cyclic, we cannot expect this theorem to hold for $p\neq 0$. Examples 1, 2 and 5 of Section 3 will show that for $p\neq 0$, a point above a simple point of the branch locus need not be simple. Example 5 will also show that even if the branch locus B for $P\rightarrow P^*$ has a simple point at P, still the image B^* of B may split at P^*. If we have a sequence $P\rightarrow P^*_1\rightarrow P^*_2\rightarrow\cdots\rightarrow P^*_q$ such that the branch locus B for $P\rightarrow P^*_q$ has a simple point at P, B may acquire a singularity at P^*_1, and hence a bad singularity at P^*_2, and a worse one at P^*_3 etc.; thus making the groups $G(P^*_i:P^*_{i-1})$ more and more complex. This "local splitting of a simple branch variety by itself" is the real reason behind the failure of Theorem 3 and 4 for $p\neq 0$, and which in turn explains the failure for nonzero characteristic fields of Jung's classical method of uniformization (see Section 4).

3. Some examples.

In the following examples we shall be dealing with a normal surface $\bar{F}$ in a projective space S_3 over algebraically closed ground field k of characteristic $p\neq 0$. $\bar{F}$ will be given by its defining homogeneous polynomial F in the homogeneous coordinates X, Y, Z, T in S_3. We shall project $\bar{F}$ from the point $X=Y=T=0$ onto the (X,Y,T) plane. By $\bar{K}$ and K we shall denote the function fields of the (X,Y,T) plane and of $\bar{F}$ respectively. By D we shall denote the branch curve of this projection on the (X,Y,T) plane. We shall omit some details in the proofs which consist of straightforward computations. For instance: since k is algebraically closed, $F=\partial F/\partial Z=\partial F/\partial X=\partial F/\partial Y=\partial F/\partial T=0$ will give the singular manifold of F (Theorem 7 of [12]). In Examples 1 to 5, F will have only a finite number of singular points, from which will follow the irreducibility of F and hence of $\bar{F}$. Consequently in Examples 1 to 5, we shall have that the singularities of F are exactly the singularities of $\bar{F}$ and the finitenes of their number will in turn imply the (arithmetic) normality of $\bar{F}$ (Theorem 3 of [8]). We shall find D by computing the Z-discriminant of F. Also we shall use well-known facts from ramification theory of real discrete valuations, for which we refer to [3] and [5].

Example 1.

$$F=\begin{cases} Z^p-X^{p-1}Z-Y^{p-1}T & \text{if } p\neq 2\\ Z^2-XZ-Y^2 & \text{if } p=2.\end{cases}$$

$\bar{F}:F=0$ is a normal surface with $X=Y=Z=0$ as the only singular point. $X=0$ is the equation of the branch curve D. *Thus* $X=Y=0$ *is a simple* *point of D and above it lies the singular point $X = Y = Z = 0$* of $\bar{F}$. $\bar{K}/K$ is a cyclic extension of degree p. Consider the real discrete valuation v_X of K/k given by the Y-axis: $X = 0$, in the (X, Y)-plane. For $p \neq 2$, reducing F modulo v_X we get: $Z^p - Y^{p-1} = 0$. This is an irreducible pure inseparable equation over the residue field $k(Y)$ of v_X. Hence v_X has a unique extension $\bar{v}_X$ to $\bar{K}$. We have: $i(\bar{v}_X : v_X) = p$ and $\bar{r}(\bar{v}_X : v_X) = 1$. For $p = 2$, writing Z^* for $Z + Y$ we get: $F = Z^{*2} + XZ^* + XY$. Again v_X has a unique $\bar{v}_X$ to $\bar{K}$, but now $\bar{r}(\bar{v}_X : v_X) = 2$ and $i(\bar{v}_X : v_X) = 1$ since $\bar{v}_X(Z^*) = \frac{1}{2}v_X(X)$.

Example 2.

$$F = Z^p - X^{p-1}Z - XY^{p-1}.$$

$\bar{F} : F = 0$ is a normal surface with $X = Y = Z = 0$ as the only singuarity. $X = 0$ is the equation of the branch curve D. Again $X = Y = 0$ *is a simple point of D and above it lies the singular point $X = Y = Z = 0$ of $\bar{F}$*. Again $\bar{K}/K$ is cyclic of degree p. Let v_X be the valuation of K/k given by $X = 0$ and let $\bar{v}_X$ be an extension of v_X to $\bar{K}$. Then $\bar{v}_X(Z) = (1/p)v_X(X)$. Hence $\bar{v}_X$ is unique, $\bar{r}(\bar{v}_X : v_X) = p$, and $i(\bar{v}_X : v_X) = 1$. (Note that Examples 1 and 2 coincide for $p = 2$).

Example 3. $p = 2$.

$$F = Z^5 + XZ^2T^2 + YZT^3 + X^4T.$$

$\bar{F} : F = 0$ is a normal surface with $X = Y = Z = 0$ and $X = T = Z = 0$ as the only singular points. Components of the branch curve D through the point $Q : X = Y = 0$ are $X = 0$ and $X^6 + Y = 0$. *So Q is an ordinary double point of D.*

The point $P : X = Y = Z = 0$ is the only point of $\bar{F}$ above Q. Let $\bar{R}$ and R be the completions of the local rings of P and Q respectively, and let $\bar{E}$ and E be the respective quotient fields of $\bar{R}$ and E. It is clear that $R = k[[X, Y]], E = k((X, Y))$, and that

$$F(Z) = Z^5 + XZ^2 + YZ + X^4$$

is the minimal polynomials of a primitive element of $\bar{E}$ over E. Let E^* be the least normal extension of E containing $\bar{E}$, and let G be the Galois group of E^* over E.

We shall show that G is insolvable.

Since E^*/E is the root field of $f(Z) = 0$ we can consider G as a subgroup of the permutation group P_5 on five symbols. Let n be the order of G.

Consider the discrete valuation v_X of $\bar{E}$ given by the prime element X of R. Reducing the equation $f(Z)=0$ modulo v_X we get:

$$Z^5+YZ=Z(Z^4+Y)=0.$$

Therefore there exists a valuation of E^* above v_X with inseparability index divisible by 4.

Similarly reducing $f(Z)=0$ module the valuation v_{Y-X^6} of E given by $Y-X^6$ we get:

$$Z^5+XZ^2+X^6Z+X^4=(Z^2+X^3)(Z^3+X^3Z+X)=0.$$

Therefore there exists a valuation of E^* above v_{X-Y^6} with residue degree divisible by 3.

Hence n is divisible by 3, 4 and 5; i. e., n is divisible by 60. Since P_5 is of order 120 and since the alternating group A_5 is the only subgroup of P_5 of order 60; it follows that G is either A_5 or P_5. In either case G is insolvable.

Example 4. $p=2$.

$$F=Z^5+XZ^4+YZ^3T+X^4Z+XY^2T.$$

Again $\bar{F}:F=0$ is a normal surface with $X=Y=Z=0$ and $X=T=Z=0$ as the only singular points. But now $X=0$ is the only component of D through the point $Q:X=Y=Z=0$. *So Q is a simple point of D.* Also $P:X=Y=Z=0$ is the only point of F above Q. Let R, $\bar{R}$, E, $\bar{E}$, E^* and G be as Example 3. Now $f(Z)=Z^5+XZ^4+YZ^3+X^4Z+XY^2$ is the minimal polynomial of a primitive element of $\bar{E}/E$. Again consider G as a subgroup of the permutation group P_5 on five symbols.

Reducing $f(Z)=0$ module the valuation v_X of E given by the prime element X of R we get:

$$Z^3(Z^2+Y)=0.$$

Therefore there exists a valuation v_1 of $\bar{E}$ above v_X with inseparability index divisible by 2. Also we have that there is an extension v_2 of v_X to $\bar{E}$ such that $v_2(z)>0$ where z is a root of $f(Z)=0$. It is clear that Yz^3 and XY^2 must be the minimum value terms in $f(z)$ for v_2, i. e., we have $v_2(Yz^3)=v_2(XY^2)$. Since $v_2(Y)=v_X(Y)=0$ and $v_2(X)=v_X(X)=1$; it follows that $v_2(z)=\frac{1}{3}$, and hence we have that $\bar{r}(v_2:v_X)\geqq 3$. Now $v_1=v_2$ implies that $r(v_1:v_X)\geqq 6>[\bar{E}:E]$ which is a contradiction. Hence $v_1\neq v_2$. Therefore $i(v_1:v_X)+\bar{r}(v_2:v_X)\leqq 5=[\bar{E}:E]$ i. e., $\bar{r}(v_2:v_X)=3$. Therefore every extension v^* of v_X to E^* is such that $i(v^*:v_X)$ is divisible by 2 and $\bar{r}(v^*:v_X)$ is divisible by 3.

Therefore the order n of G is divisible by 2, 3, 5 and hence by 30. Hence G is either (i) P_5, or (ii) A_5, or (iii) a subgroup of P_5 of order 30. In cases (i) and (ii) G is insolvable. We shall show that case (iii) cannot arise by proving that *P_5 has no subgroup of order* 30. Suppose then, if possible, that H is a subgroup of P_5 of order 30. Let $H = H_1, H_2, H_3, H_4$ be the right cosets of H, and for $u \varepsilon P_5$ let

$$(tu)(H_1, H_2, H_3, H_4) = (H_1u, H_2u, H_3u, H_4u).$$

Then t is a homomorphism of P_5 into the group of permutations of H_1, H_2, H_3, H_4. Since $u \varepsilon H_i$ implies $H_1u = H_i$, $4 \leqq$ order of $t(P_5) \leqq 24$. Hence $t^{-1}(0)$ is a normal subgroup of P_5 of order m, where $(120)/(24) = 5 \leqq m \leqq 30 = 120/4$. A contradiction since A_5 is the only normal subgroup of P_5. *Therefore G is insolvable.*

Example 5.

$$F = Z^{p+1} + Y^{p-1}ZT + X^{p+1}.$$

The surface $\bar{F} : F = 0$ is normal and has $X = Y = Z = 0$ and $X = T = Z = 0$ as the only singular point. *The point $Q : X = Y = 0$ is a simple point of the branch curve $D : X = 0$.* Again in the terminology of Example 3,

$$f(Z) = Z^{p+1} + Y^{p-1}Z + X^{p+1},$$

is the minimal polynomial of a primitive element of $\bar{E}/E$. Consider the Galois group G of E^*/E, as a subgroup of the permutation group P_{p+1} on the roots of $f(Z) = 0$. Modulo the valuation v_X of E given by the prime element X in R the equation $f(Z) = 0$ becomes $Z(Z^p + Y^{p-1}) = 0$. Hence there exists a valuation of E^* above v_X with inseparability index divisible by p. Since the order n of G must then be divisible by p and since p is prime, G must contain a p-cycle. Thus G is a subgroup of P_{p+1}, G contains a p-cycle, and the order of G is divisible by $p(p+1)$. For $p \geqq 5$ such a group G is very large and quite complicated.

Observe that *the point $P = X = Y = Z = 0$ is above the simple point $Q : X = Y = 0$ of the branch curve $D : X = 0$*; and still *P is singular. D splits in passing from Q to P* as is shown by the equation: $Z(Z^p + Y^{p-1}) = 0$.

Example 6. $p = 5$.

$$F = Z^5 + YZ^3 + YZ^2 + X.$$

$\bar{F} : F = 0$ is nonsingular at finite distance. *$Q : X : Y = 0$ is an ordinary double point of the local branch curve $XY = 0$.* The only point of $\bar{F}$ above Q is

$P: X = Y = Z = 0$. *P is a simple point of* $\bar{F}$ *but still the local Galois group* G *of* P *over* Q (notation of Example 3) *is insolvable.*

Modulo the valuation v_X of E given by X, we have

$$Z^5 + YZ^3 + YZ^2 = Z^2(Z^3 + YZ + Y) = 0.$$

Hence there exists a valuation v_1 of $\bar{E}$ over v_X with separable residue degree over v_X divisible by 3. Also there is a valuation v_2 of $\bar{E}$ above v_X for which $v_2(z) > 0$ where z is a root of $F(Z) = 0$. Since for v_2, Yz^2 and X are the minimum value terms of $F(z)$; it follows that $v_2(z) = (\frac{1}{2})v_X(X) = \frac{1}{2}$. As in Example 4 we conclude that: $v_1 \neq v_2$, $\bar{r}(v_2 : v_X) = 2$, and $g(v_1 : v_X) = 3$. (One extension with $g = 3$ and one extension with $\bar{r} = 2$ makes up $3 + 2 = 5 = [\bar{E} : E]$ and accounts for all the $\bar{E}$-extensions of v_X). Hence the order of G is divisible by $2 \times 3 \times 5 = 30$. The insolvability of G now follows as in Example 4.

Remark. In the above examples, let Q and P be the points $X = Y = 0$ and $X = Y = Z = 0$ of the (X, Y, T) plane and of $\bar{F}$ respectively. Observe that in all the examples: $r(P:Q) = \bar{r}(P:Q) = [\bar{K}:K] = [\bar{E}:E]$. Now let K^* be the least Galois extension of K containing $\bar{K}$, and let P^* be a point corresponding to P on a derived normal model F^* of the (X, Y, T) plane in K^*. By Proposition 1, D is the branch curve also for the rational transformation from F^* onto the (X, Y, T) plane. In Example 3, Q is an ordinary double point of D and in Examples 4 and 5, Q is a simple point of D. In Examples 3 and 4, $G(P^*:Q)$ is insolvable. In Example 5, since D splits in passing from Q to P, D splits *a fortiori* in passing from Q to P^*. In Examples 1 and 2, $\bar{K} = K^*$ and hence $P = P^*$.

4. Jung's method of uniformization. Let S be an algebraic surface in an n-dimensional projective space over an algebraically closed ground field k and let K be the function field of S. The classical problem of local uniformization on algebraic surfaces deals with the case when k is the field of complex numbers. It roughly asserts the following: Let P be a point on S. Then some complete neighborhood of P on S can be covered by a finite number of regions $R_1, R_2, \cdots, R_t$ such that in each R_i the coordinates $x_1, x_2, \cdots, x_n$ of the points in R_i can be expressed as convergent power series in two variables u, v; this representation being unramified in a certain sense. This problem was solved by Jung[2] by using a very fruitful method. Zariski generalized this problem when k is an abstract field, stating it in terms of valuation theory. If we use the concept of normal varieties due to Zariski, then the basic ideas behind Jung's proof become strikingly simple and are as described below.

Let V be a normal model of the function field K/k in a projective space S_r or dimension r. Project V onto a plane V', from an S_{r-3} which does not meet V. Let K'/k be the function field of V'. Let T be the rational transformation from V onto V', thus defined. Since T and T^{-1} are free from fundamental points, V is a derived normal model of V' in K. By Theorem 2, the branch locus D on V is a curve. Let Q be a singular point of D. Apply to V' a locally quadratic transformation t, with center Q, getting a new surface V'_1. To Q there will correspond a fundamental line L_1 on V'_1. We have that there is a rational transformation tT of V onto V'_1 whose inverse is finitely valued at each point of V'_1. Let V_1 denote a derived normal model of V'_1 in the function field K/k. Let T_1 be the rational transformation of V_1 onto V'_1 (defined in a natural fashion). Then V_1, V'_1 and T_1 are in the same relationship to each other as V, V' and T. Also V_1 has no singular points. The branch curve D_1 of T_1^{-1} is either the proper transform of D, or it is this transform together with the line L_1. After a finite number of steps it is possible to get V_1, V'_1 and T_1 such that the branch curve D_1 on V'_1 has only ordinary double points. We may assume that already V, V', and T have this property, i. e., that D has only ordinary double points.

Let Q be a point of V' and let P be a point of V corresponding to Q. We have proved in Section 2 that: (a) If $Q \notin D$, then P is a simple point of V. (b) If Q is a simple point of D, then $G(P/Q)$ is cyclic and again P is a simple point. (c) If Q is an ordinary double point of D, then $G(P/Q)$ is the direct product of two cyclic groups.[12] So it is only necessary to consider what happens "above" an ordinary double point Q of the branch curve D. Let P be a point of V above Q. The simple structure of $G(P/Q)$ in this case enables Jung to resolve the singularity of V at Q in the classical case when k is the field of complex numbers, by topological considerations.

Now supposing that k is of characteristic $p \neq 0$, we tried to generalize Jung's method. But unfortunately in this case of modular ground fields as was shown in Section 3, the statements (b) and (c) are not valid and the groups mentioned there may even be insolvable. If these local Galois groups were at least solvable, that would have been of some use for the purposes of uniformization.

HARVARD UNIVERSITY.

BIBLIOGRAPHY.

[1] S. Abhyankar, "Local uniformization on algebraic surfaces over ground fields of characteristic $p \neq 0$," forthcoming in the *Annals of Mathematics*.

[2] C. Chevalley, "Some properties of ideals in rings of power series," *Transactions of the American Mathematical Society*, vol. 55 (1944), pp. 68-84.

[3] M. Deuring, "Verzweigungstheorie bewerteter Körper," *Mathematische Annalen*, vol. 105 (1931), pp. 277-307.

[4] H. W. E. Jung, "Darstellung der Funktionen eines algebraischen Körpers zweier unabhängigen Veränderlichen in der Umgebung einer Stelle," *Journal Für Mathematik*, vol. 133 (1908), pp. 289-314.

[5] W. Krull, "Galoissche Theorie bewerteter Körper," *Sitzungbereichte der Bayerschen Akad. der Wissenschaften*, München (1930).

[6] ———, "Dimensionstheorie in Stellenringe," *Journal Für Mathematik*, vol. 179 (1938), pp. 204-226.

[7] ———, "Der allgemeine Diskriminantensatz," *Mathematische Zeitschrift*, vol. 45 (1939), pp. 1-19.

[8] A. Seidenberg, "The hyperplane sections of normal varieties," *Transactions of the American Mathematical Society*, vol. 69 (1950), pp. 357-383.

[9] B. L. van der Waerden, *Modern Algebra*, Vol. II, New York, 1950.

[10] R. J. Walker, "Reduction of the singularities of an algebraic surface," *Annals of Mathematics*, vol. 36 (1935), pp. 336-365.

[11] O. Zariski, "Foundations of a general theory of birational correspondences," *Transactions of the American Mathematical Society*, vol. 53 (1943), pp. 490-542.

[12] ———, "The concept of a simple point of an abstract algebraic variety," *ibid.*, vol. 62 (1947), pp. 1-52.

[13] ———, *Theory and Applications of Holomorphic Functions on Algebraic Varities over Arbitrary Ground Fields*, New York, 1951.

[14] ———, "Abstract Algebraic Geometry, Vol. I," forthcoming in the *American Mathematical Society* Publications.

A NOTE ON TWO DIMENSIONAL DIVISION RING EXTENSIONS.*[1]

By N. JACOBSON.

A division subring Γ of a division ring Δ is said to be *Galois in* Δ (and Δ is *Galois over* Γ) if Γ is the set of invariants (or fixed points) of a group of automorphisms acting in Δ. The two dimensional Galois extensions of a division ring have been determined by Dieudonné.[2] In this note we shall show that if Γ is a division ring of characteristic $\neq 2$ which is finite dimensional over its center Ψ and Δ contains Γ and has left dimensionality $[\Delta:\Gamma]_L = 2$ then Δ is Galois over Γ. On the other hand, we shall construct a class of examples where $[\Delta:\Gamma]_L = 2 = [\Delta:\Gamma]_R$, Γ of characteristic $\neq 2$ and Δ is not Galois over Γ.

We prove first a general theorem on dimensionalities as follows.

THEOREM 1. *Let* Δ *be a division ring,* Γ *a division subring,* Φ *and* Ψ *the centers of* Δ *and* Γ *respectively.* (1) *If* $[\Delta:\Gamma]_L < \infty$ *and* $[\Gamma:\Psi] < \infty$ *then* $[\Delta:\Phi] < \infty$. (2) *If* $[\Delta:\Phi] < \infty$ *then* $[\Gamma:\Psi] \leqq [\Delta:\Phi]$, *the equality holding if and only if the division subring* $\Phi(\Gamma)$ *generated by* Φ *and* Γ *coincides with* Δ.

Proof. (1) Our assumptions here imply that $[\Delta:\Psi]_L < \infty$. The result will therefore follow from the fact, which we proceed to prove, that if a division ring Δ is of finite left dimensionality over a subfield Ψ then Δ is of finite dimensionality over its center. Let $\delta_1, \delta_2, \cdots, \delta_n$ be elements of Δ which are linearly independent over Φ. Then it is well known that the right multiplications $\delta_{1R}, \delta_{2R}, \cdots, \delta_{nR}$ are linearly independent over the division ring Δ_L of left multiplications. Hence δ_{iR} are linearly independent over Ψ_L the subfield of Δ_L determined by Ψ. On the other hand, the ring $\mathfrak{L}$ of linear transformations of the left vector space Δ over Ψ contains Ψ_L since Ψ is commutative. Moreover, $[\mathfrak{L}:\Psi_L] \leqq [\Delta:\Psi]_L^2$. Since the $\delta_{iR} \varepsilon \mathfrak{L}$ this implies that $n \leqq [\Delta:\Psi]_L^2$. Hence $[\Delta:\Phi] \leqq [\Delta:\Psi]_L^2$. (2) If $x_1, x_2, \cdots, x_r$ are

---

* Received January 4, 1955.

[1] This research was supported in part by a grant from the National Science Foundation.

[2] [4]. Cf. also [1].

elements of a ring then we define $[x_1x_2\cdots x_r] = \sum_P \pm x_{i_1}x_{i_2}\cdots x_{i_r}$ where the summation is over all the permutations $(i_1, i_2, \cdots, i_r)$ of $(1, 2, \cdots, r)$ and the sign is + or — according as the permutation is even or odd. It has been shown by Amitsur and Levitzki ([2]) that if Φ_n is the ring of $n \times n$ matrices over a field Φ then $[x_1x_2\cdots x_{2n}] = 0$ for all $x_i \varepsilon \Phi_n$ but there exist $x_1, x_2, \cdots, x_{2n+1}$ in Φ_n such that $[x_1x_2\cdots x_{2n+1}] \neq 0$. This implies by the usual base field extension argument that if Δ is a division ring which is n^2 dimensional[3] over its center then $[x_1x_2\cdots x_{2n}] \equiv 0$ in Δ but $[x_1x_2\cdots x_{2n+1}] \not\equiv 0$ in Δ. Also it is a known result of Kaplansky's ([7]) that if $[\Delta:\Phi] = \infty$ then Δ satisfies no polynomial identity with coefficients in Φ. It is clear from these results that $[\Delta:\Phi] < \infty$ implies $[\Gamma:\Psi] \leqq [\Delta:\Phi]$. Now consider $\Phi(\Gamma)$. This is a subspace of Δ over Φ; hence it is finite dimensional. Let $\gamma_1, \gamma_2, \cdots, \gamma_r$ be a maximal set of elements of Γ which are linearly independent over Φ and let V be the Φ-space spanned by these. Evidently $\Gamma \subseteq V$. Hence V is a subring of Δ. Since V is finite dimensional over Φ it is a division subring. Thus $V = \Phi(\Gamma)$. Suppose $[\Gamma:\Psi] = m^2$, then $[\gamma_{i_1}\gamma_{i_2}\cdots\gamma_{i_{2m}}] = 0$ for all choices of $i_j = 1, 2, \cdots, r$. Hence $[x_1x_2\cdots x_{2m}] \equiv 0$ in $\Phi(\Gamma) = V$ so that if P is the center of $\Phi(\Gamma)$, then $[\Phi(\Gamma):\mathrm{P}] \leqq m^2$. On the other hand, if $[x_1x_2\cdots x_{2k}] \equiv 0$ in $\Phi(\Gamma)$ then the same holds in Γ. It follows that $[\Gamma:\Psi] \leqq [\Phi(\Gamma):\mathrm{P}]$. Hence $[\Gamma:\Psi] = [\Phi(\Gamma):\mathrm{P}] = m^2$. Now if $[\Gamma:\Psi] = [\Delta:\Phi] = n^2$ then $[\Phi(\Gamma):\mathrm{P}] = n^2$ and, since $\mathrm{P} \supseteq \Phi$, $[\Phi(\Gamma):\Phi] = n^2$. Thus $\Phi(\Gamma) = \Delta$. The converse is clear by the same argument.

THEOREM 2. *Let Γ be finite dimensional over its center Ψ and let Δ be a division ring containing Γ and such that $[\Delta:\Gamma]_L = 2$. Then if the characteristic is $\neq 2$, Δ is Galois over Γ.*

Proof. Theorem 1 shows that $[\Delta:\Phi] < \infty$ for Φ the center of Δ. We distinguish two cases; I. $\Gamma \supset \Phi$. Let $\mathfrak{C}(\Gamma)$ denote the centralizer of Γ in Δ. Then it is well known that $\mathfrak{C}(\mathfrak{C}(\Gamma)) = \Gamma$ so that Γ is the set of invariants of the inner automorphisms determined by the non-zero elements of $\mathfrak{C}(\Gamma)$. II. $\Gamma \not\supset \Phi$. Let $t \varepsilon \Phi, \not\varepsilon \Gamma$. Then $\Delta = \Gamma(t)$ and $t^2 = \gamma_1 t + \gamma_2$, $\gamma_i \varepsilon \Gamma$. Since t and t^2 commute with every $\gamma \varepsilon \Gamma$ this gives $(\gamma\gamma_1)t + \gamma\gamma_2 = (\gamma_1\gamma)t + \gamma_2\gamma$. Hence $\gamma\gamma_i = \gamma_i\gamma$ and so that γ_i are in the center Ψ of Γ. Since $\Delta = \Gamma(t)$ where $t \varepsilon \Phi$ it is clear that $\Psi = \Gamma \cap \Phi$. Hence the $\gamma_i \varepsilon \Phi$. We may replace t by $u = t - \frac{1}{2}\gamma_1$ and obtain $u^2 = \delta \varepsilon \Phi$. For $\alpha, \beta \varepsilon \Gamma$, the mapping $\alpha + \beta u \rightarrow \alpha - \beta u$ is an automorphism in Δ whose set of invariants is Γ.[4]

---

[3] It is well known that if Δ is finite over its center then its dimensionality over the center is a square.

[4] I am indebted to the referee for the following remark: In part II of the above

In the remainder of this note we shall construct some examples of two dimensional non-Galois extensions. These will be given as certain Clifford algebras of infinite dimensionality. We shall therefore begin by indicating that the well-known results on finite dimensional Clifford algebras can be carried over to the infinite case.

Let M be a vector space over a field Φ and let $Q(x)$ be a quadratic form on M. We denote the associated scalar product $Q(x+y) - Q(x) - Q(y)$ by (x, y). This is symmetric and if the characteristic is two then (x, y) is alternate: $(x, x) = 0$. If N is a subspace we define $N^{\perp} = \{y \mid (x, y) = 0, x \varepsilon N\}$. N is isotropic if $N \cap N^{\perp} \neq \{0\}$, totally isotropic if $N \subseteq N^{\perp}$. The scalar product (and the quadratic form) is non-degenerate if $M^{\perp} \neq \{0\}$, totally regular if there exists no isotropic $((x, x) = 0)$ $x \neq 0$.

Let $\mathfrak{F}$ be the free algebra $\mathfrak{F} = \Phi 1 \oplus M \oplus (M \otimes M) \oplus (M \otimes M \otimes M) \oplus \cdots$ based on the space M and let $\mathfrak{B}$ be the ideal in F generated by $x \otimes x - Q(x)1$. Then $\Delta = \mathfrak{F}/\mathfrak{B}$ is called the *Clifford algebra* of M relative to Q.

It is easy to see that the natural homomorphism of M into Δ is an isomorphism ([3], p. 39). One may therefore consider M as a subspace of Δ. If N is a subspace of M, then the subalgebra of Δ generated by 1 and N is isomorphic, and so may be identified with the Clifford algebra $\Delta(N, Q)$ of N relative to the contraction of Q.[5]

From now on we assume that (x, y) is non-degenerate. Then it is well known that if M is even (finite) dimensional then Δ is central simple and if M is odd dimensional then Δ is a Kronecker product of a central simple algebra and a two dimensional semi-simple algebra. The structure of Δ can be described more precisely by using the following

LEMMA. *Let F be a non-isotropic subspace of $2\mu < \infty$ dimensions. Then the Clifford algebra $\Delta(M, Q)$ of M relative to Q is a Kronecker product $\Delta(F, Q) \otimes \Delta(F^{\perp}, (-1)^{\mu(2\mu-1)}\delta Q)$ of the Clifford algebra of F and Q and that of $F^{\perp}$ and $(-1)^{\mu(2\mu-1)}\delta Q$ where δ is the discriminant of (x, y) in F.*

Proof. Char. $\neq 2$. We choose an orthogonal basis $(u_1, u_2, \cdots, u_{2\mu})$ for F and set $(u_i, u_i) = \delta_i \neq 1$. Set $c = 2^{2\mu} u_1 u_2 \cdots u_{2\mu}$. Since $u_i u_j = -u_j u_i$, for $i \neq j$, $u_i^2 = \frac{1}{2}\delta_i 1$, $c^2 = (-1)^{\mu(2\mu-1)}\delta$ where $\delta = \Pi \delta_i$ is the discriminant of (x, y). It is known, that for finite dimensional non-isotropic subspaces

---

proof, no use is made of dimensionality relations. Hence the argument shows that if $[\Delta : \Gamma]_L = 2$ for arbitrary Γ, and $\Gamma \not\supseteq \Phi$, then Δ is the Kronecker product over Ψ of Γ and a quadratic extension of Ψ.

[5] It suffices to prove this for finite dimensional N. The arguments given in [3] suffice in this case. If N is even dimensional and non-isotropic then the lemma given below can also be used. An easy reduction can be made to this special case.

$F, M = F \oplus F^{\perp}$. If $x \varepsilon F^{\perp}$, $xc = cx$ and $(cx)^2 = Q(x)(-1)^{\mu(2\mu-1)}\delta$. It follows easily that the subalgebra generated by the elements cx is the Clifford algebra $\Delta(F^{\perp}, (-1)^{\mu(2\mu-1)}\delta Q)$ and

$$\Delta(M, Q) = \Delta(F, Q) \otimes \Delta(F^{\perp}, (-1)^{\mu(2\mu-1)}\delta Q).$$

Char. $= 2$. Here we can choose a symplectic basis for F:

$$(u_1, v_1, u_2, v_2, \cdots, u_\mu, v_\mu)$$

with $(u_i, v_i) = 1 = (v_i, u_i)$ and all other products 0. The discriminant of (x, y) relative to this basis is 1. If $x \varepsilon F$, $y \varepsilon F^{\perp}$ then $xy = yx$. It follows that $\Delta(M, Q) = \Delta(F, Q) \otimes \Delta(F^{\perp}, Q)$.

This lemma shows that if M is of even dimensionality then Δ is a Kronecker product of generalized quaternion algebras and if M is odd dimensional then Δ is a Kronecker product of generalized quaternion algebras and a two dimensional semi-simple algebra. We now prove

THEOREM 3. *Let M be an infinite dimensional vector space with a non-degenerate quadratic form Q. Then the Clifford algebra $\Delta(M, Q)$ is central simple. Moreover, if $\dim M = \aleph_0$ then $\Delta(M, Q)$ is a Kronecker product of a countable number of generalized quaternion algebras and any such product can be obtained from a suitable Clifford algebra.*

Proof. The first statement will follow by showing that any finite subset S of $\Delta = \Delta(M, Q)$ is contained in a finite dimensional central simple subalgebra containing 1. By using a basis one sees that S is contained in the subalgebra generated by 1 and a finite dimensional subspace F. Since any finite dimensional subspace can be imbedded in a non-isotropic one, we may assume F non-isotropic. Also it is easy to see (using $M = F \oplus F^{\perp}$) that we may assume $\dim F = 2\mu$. Then the subalgebra $\Delta(F, Q)$ is central simple. Now suppose $\dim M = \aleph_0$. If char $\neq 2$ we denote the generalized quaternion algebra with basis $(1, i, j, ij)$ such that $i^2 = \alpha 1, j^2 = \beta 1$, $ij = -ji$ by (α, β) and if char. $= 2$ we denote the generalized quaternion algebra with basis $(1, i, j, ij)$ such that $i^2 = \alpha 1$, $j^2 = \beta 1$, $ij + ji = 1$ by $[\alpha, \beta]$. If char. $\neq 2$ we choose an orthogonal basis $(u_1, v_1, u_2, v_2, \cdots)$ with $Q(u_i) = \alpha_i$, $Q(v_i) = \beta_i$ ([8], p. 3). Then using $(\alpha\rho^2, \beta\rho^2) = (\alpha, \beta)$ for all $\rho \neq 0$ and the lemma, we obtain

$$\Delta = (\alpha_1, \beta_1) \otimes (-\alpha_1\beta_1\alpha_2, -\alpha_1\beta_1\beta_2) \otimes (-\alpha_2\beta_2\alpha_3, -\alpha_2\beta_2\beta_3)$$
$$\otimes \cdots \otimes (-\alpha_\nu\beta_\nu\alpha_{\nu+1}, -\alpha_\nu\beta_\nu\beta_{\nu+1}) \otimes \cdots.$$

If char. $= 2$ we choose a symplectic basis $(u_1, v_1, u_2, v_2, \cdots)$ with $(u_i, v_i) = 1 = (v_i, u_i)$ ([8], p. 4). If $Q(u_i) = \alpha_i$, $Q(v_i) = \beta_i$ then the above lemma gives the formula

$$\Delta = [\alpha_1, \beta_1] \otimes [\alpha_2, \beta_2] \otimes [\alpha_3, \beta_3] \otimes \cdots.$$

The last statement of the theorem is clear from the above formulas.

It has been shown by Koethe ([9]) that there exist infinite Kronecker products of quaternion algebras which are division algebras. We shall sketch an alternative proof. Let Γ be a given division ring with center Ψ. Then the polynomial domain $\Gamma[\lambda]$ is a principal ideal domain and so it can be imbedded in a quotient field $\Gamma(\lambda)$. The center of $\Gamma(\lambda)$ is $\Psi(\lambda)$. If char. $\neq 2$, we let S be the automorphism in $\Gamma(\lambda)$ which is the identitly on Γ and sends λ into $-\lambda$ and we form the semi-linear polynomial domain $\Gamma(\lambda)[t, S]$ of polynomials in t with coefficients in $\Gamma(\lambda)$ such that $ta = (aS)t$, $a \in \Gamma(\lambda)$ ([5], p. 29). This is a principal ideal domain and can be imbedded in a quotient field Δ. The center Φ of Δ is $\Psi(\lambda^2, t^2)$ and $\Delta = \Gamma(\lambda^2, t^2) \otimes_\Phi (\lambda^2, t^2)$ where $\Gamma(\lambda^2, t^2)$ is the division subring generated by Γ, λ^2, t^2. We can begin our construction with a field $\Gamma = \Phi_0$ and repeat this an infinite number of times to obtain a division algebra with center $\Phi(\lambda_1{}^2, t_1{}^2, \lambda_2{}^2, t_2{}^2, \cdots)$ which has the form $(\lambda_1{}^2, t_1{}^2) \otimes (\lambda_2{}^2, t_2{}^2) \otimes \cdots$ over this center. If char. $= 2$ we modify the construction by replacing the automorphism S by the derivation D in $\Gamma(\lambda)$ such that $\Gamma D = \{0\}$ and $\lambda D = 1$. Then we consider the differential polynomial domain $\Gamma(\lambda)[t, D]$ with $ta = at + aD$. In the end we obtain the center $\Phi_0(\lambda_1{}^2, t_1{}^2, \lambda_2{}^2, t_2{}^2, \cdots)$ and the division algebra

$$[\lambda_1{}^2, t_1{}^2] \otimes [\lambda_2{}^2, t_2{}^2] \otimes \cdots$$

over this center.

We now suppose that $\Delta = \Delta(M, Q)$ is a Clifford algebra of a space M relative to a quadratic form Q and that Δ is a division ring. Let H be a hyperplane in M, so that $M = H \oplus \Phi u$ where $u \in M$ and Φ is the base field. Let $\Gamma = \Delta(H, Q)$ be the subalgebra generated by 1 and H. Since the elements of Δ are algebraic over Φ, Γ is a division ring. We have the relation $ux = -xu + (x, u)1$ for $x \in H$. Hence if $a \in \Gamma$ then we have a relation of the form $ua = (aI)u + aD$ where aI and aD are elements in Γ. Since $u \notin \Gamma$ the elements aI and aD are uniquely determined by a. One verifies that I is an isomorphism of Γ into itself and D is an I-derivation: $(ab)D = (aI)(bD) + (aD)b$. Since the image under I contains H it is clear that I is an automorphism. The relations $ua = (aI)u + (aD)$, $u^2 = Q(u)1$ show that

$(1, u)$ is both a left basis and a right basis for Δ over Γ. Hence $[\Delta : \Gamma]_L = 2 = [\Delta : \Gamma]_R$. We now prove

THEOREM 4. *Let* $\Delta = \Delta(M, Q)$ *be a Clifford division algebra determined by a vector space* M *relative to a quadratic form* Q *and let* $\Gamma = \Delta(H, Q)$ *be the subalgebra generated by* 1 *and a hyperplane* H. *Assume char.* $\neq 2$. *Then* Δ *is Galois over* Γ *if and ony if* $H^{\perp} \neq \{0\}$.

Proof. Since Δ is a division ring, M contains no isotropic vectors. Hence $H \cap H^{\perp} \neq \{0\}$. If $H^{\perp} \neq \{0\}$ we choose u in this space. Then we have $ua = (aI)u$ and we have an automorphism mapping every $a \varepsilon \Gamma$ into itself and mapping u into $-u$. The set of invariants relatively to this automorphism is Γ. Hence Γ is Galois in Δ. Conversely, assume Δ is Galois over Γ. Then it is known that $\Delta = \Gamma(t)$ where t is an element such that the inner automorphism determined by t maps Γ into itself. Write $t = cu + d, c, d \varepsilon \Gamma$, $c \neq 0$. We may replace t by $v = c^{-1}t = u + e$, $e = c^{-1}d$. If $a \varepsilon \Gamma$ the condition $va = a_1 v$, $a_1 \varepsilon \Gamma$ gives $a_1 = aI$ and $aD = (aI)e - ea$. In particular, if $x \varepsilon H$, then $(x, u)1 = -(xe + ex)$. If M is finite dimensional then it is clear that $H^{\perp} \neq \{0\}$. Hence we may assume M infinite dimensional. Then we can find an even dimensional subspace F in H such that $e \varepsilon \Delta(F, G)$ the subalgebra generated by 1 and F. Since Q is non-degenerate in F we can find a $u' \varepsilon F$ such that $(x, u') = (x, u)$ for all $x \varepsilon F$. Hence if $v' = u' + e$ then $xv' + v'x = 0, x \varepsilon F$. If $(u_1, v_1, \cdots, u_\mu, v_\mu)$ is an orthogonal basis for F, the element $u_1 v_1 \cdots u_\mu v_\mu$ has an inverse and anti-commutes with every $x \varepsilon F$. Since $\Delta(F, Q)$ is central this implies that $v' = \rho u_1 v_1 \cdots u_\mu v_\mu$. We now imbed F in a space G in H of $(2\mu + 2)$-dimensions with orthogonal basis $(u_1, v_1, \cdots, u_{\mu+1}, v_{\mu+1})$. Since $e = \rho u_1 v_1 \cdots u_\mu v_\mu - u'$, the relation $(u_{\mu+1}, u)1 = -(u_{\mu+1}e + eu_{\mu+1})$ gives $(u_{\mu+1}, u)1 = -2\rho u_1 v_1 \cdots u_\mu v_\mu u_{\mu+1}$. Hence $\rho = 0$ and $e = -u' \varepsilon H$. Then $(x, u) = (x, u')$ for all $x \varepsilon H$ and $u - u' \neq 0$ is in $H^{\perp}$, Thus $H^{\perp} \neq \{0\}$.

If M is an $\aleph_0$ dimensional space over a field of characteristic $\neq 2$ and Q is a quadratic form whose associated scalar product is non-degenerate, then we can always find hyperplanes H in M such that $H^{\perp} = \{0\}$. This is easily seen by using an orthogonal basis. This remark and Theorem 4 show that there exist two dimensional non-Galois extensions of division rings of Char. $\neq 2$.

YALE UNIVERSITY.

REFERENCES.

[1] A. S. Amitsur, "Non-commutative cyclic fields," *Duke Mathematical Journal*, vol. 21 (1954), pp. 87-106.

[2] ——— and J. Levitzki, "Minimal identities for algebras," *Proceedings of the American Mathematical Society*, vol. 1 (1950), pp. 449-463.

[3] C. C. Chevalley, *The Algebraic Theory of Spinors*, New York, 1954.

[4] J. Dieudonné, "Les extensions quadratiques des corps non-commutatifs et leurs applications," *Acta Mathematica*, vol. 87 (1952), pp. 175-242.

[5] N. Jacobson, *The Theory of Rings*, New York, 1943.

[6] ———, "Structure of alternative and Jordan bimodules," *Osaka Mathematics Journal*, vol. 6 (1954), pp. 1-71.

[7] I. Kaplansky, "Rings with a polynomial identity," *Bulletin of the American Mathematical Society*, vol. 54 (1948), pp. 575-580.

[8] ———, "Forms in infinite-dimensional spaces," *Academie Brasiliera de Ciencias*, vol. 22 (1950), pp. 1-16.

[9] G. Koethe, "Schiefkörper unendlichen Ranges über dem Zentrum," *Mathematische Annalen*, vol. 105 (1931), pp. 15-39.

ON THE ORTHOGONALIZATION OF OPERATOR REPRESENTATIONS.*

By RICHARD V. KADISON.

---

1. **Introduction.** The main problem with which we shall be concerned is that of finding conditions under which a group representation is similar to a unitary representation and conditions for a representation of a self-adjoint algebra of operators on a Hilbert space to be similar to an adjoint preserving representation (* representation). These situations are slightly different aspects of the representation orthogonalization process with very close interconnections. With regard to the algebra situation one can phrase the main question in the following way. *Is an algebra of operators on a Hilbert space which is the isomorphic image of a C*-algebra (uniformly closed self-adjoint algebra of operators) similar to a C*-algebra in such a way that the composition of the isomorphism and the similarity is an adjoint preserving representation of the C*-algebra?* The question for group representations takes the following form. *Is every bounded representation of a group by operators on a Hilbert space similar to a unitary representation, where, by "bounded representation" we mean that there exists a constant such that each representing operator is, in norm, less than this constant?* In this form, the group question has been raised before, notably in [1], [2], [5]. The question of when a group admits a mean (or Banach Limit) is the primary consideration of these papers and then, employing the technique of [4], it is shown that for such groups all bounded representations (continuous in the strong topology) are similar to unitary representations (this is done in [4] for the infinite cyclic group and the reals). We outline this technique. A (left, right) mean on a group G is a positive linear functional on the linear space $B(G)$ of bounded continuous functions on G which takes the value 1 at the constant function 1 and which is invariant under (left, right) translations on the group. If G admits a right mean m and $g \rightarrow A_g$ is a strongly continuous representation of the group by operators on a Hilbert space, then, for each x, y in the Hilbert space $g \rightarrow (A_g x, A_g y)$ is a bounded continuous function on G and $\langle x, y\rangle = m((A_g x, A_g y))$ is an inner product on the Hilbert

---

* Received February 17, 1954; revised February 1, 1955.

space giving rise to a norm equivalent to the original norm and under which the operators A_g are unitary. It is then standard to find the similarity of the original Hilbert space which takes each A_g into a unitary operator. Continuity considerations do not enter into the question of whether all bounded representations of all groups are similar to unitary representations since it is obviously sufficient to settle this for discrete groups. With regard to the technique of means, it is well-known that many groups admit neither a left nor right mean (in the sense noted above, e. g., the free group on two generators; see [1], [2]) so that this method cannot, in itself, give the full answer.

Our concern is not with conditions on the group which imply that bounded representations are similar to unitary representations but rather with restrictions upon the representations which insure that they are similar to unitary representations. We feel that this approach gives hope of settling the full question one way or the other.

The truth of the operator algebra proposition is trivially implied by the truth of the group statement (see the proof of Theorem 7). On the other hand, a bounded representation of a group (discrete) can be extended to a representation of its L_1 algebra (as a self-adjoint algebra of operators acting by convolution on L_2). This algebra of operators need not be closed in the uniform operator topology and the representation need not be extendable to the uniform closure of this algebra; so that it would appear that the truth of the algebra result would imply the group conjecture only for those bounded representations which are extendable to the uniform closure of the L_1 algebra. (The similarity which transforms the representation of the L_1 algebra into a $*$ representation will transform the group representation into a unitary representation). However, it is possible to renorm the L_1 algebra in such a way that the completion of the resulting $*$ algebra is a C^*-algebra to which each bounded group representation is extendable, by assigning to each element of the L_1 algebra the supremum of the norms of its images in each $*$ representation. It follows from the existence of a Banach algebra norm on L_1 in which the $*$ map is isometric (viz., the L_1 norm) that this supremum is not greater than the L_1 norm (and certainly finite). Our extended group representation, being a bounded algebra representation relative to the L_1 norm on the L_1 algebra of the group, is a continuous representation of the L_1 algebra in the C^* norm just constructed. Although the representation extended to this C^*-algebra may not be an isomorphism, the kernel is a closed two-sided ideal so that the factor algebra is a C^*-algebra [8], and the induced representation on this factor C^*-algebra is an isomorphism. Thus we have the complete equivalence of the group and C^*-algebra questions. We are indebted to I. Kaplansky for bringing to our attention the known renorming device used above. It is possible that a more incisive operator algebra result would apply directly to the L_1 algebra (acting by convolution on L_2). In [6], Mackey proves the algebra result in the commutative case by direct methods (this result follows at once, as in the proof of Theorem 7, from the fact that commutative groups have means, in the sense defined above; see [5]).

The origin of the group question can be found in the classical statement which says that each representation of a finite group by (complex, real) matrices is equivalent to a representation by (unitary, orthogonal) matrices and its extension to continuous representations of compact groups [7]. The technique used in these proofs, invariant integration over the group, is almost identical with the technique of means. Using this theorem for compact groups the operator algebra result follows for finite-dimensional operator algebras (applied to the (compact) group of unitary operators in the algebra). Perhaps a more natural way of concluding the algebra result in the finite-dimensional case is thru the semi-simplicity of the image algebra. In this case the various concepts of semi-simplicity coincide so that the semi-simplicity of the original C^*-algebra, interpreted algebraically, is inherited by the image, and this, interpreted spatially, shows that this image is similar to a C^*-algebra. In the infinite-dimensional case it is not at all difficult to construct an algebra of operators which is semi-simple in all the conventional senses but not similar to a C^*-algebra. The following topological difficulty can occur: while each invariant subspace may have a complementary invariant subspace, the greatest of the angles between the given space and all possible invariant complements may tend to 0 for some sequence of invariant subspaces, thus ruling out what we may call the "topological semi-simplicity" of the algebra. For a C^*-algebra and a unitary group, the orthogonal complement of an invariant subspace is invariant. Our similarity problem for the given family of operators (group or algebra) amounts to an orthogonalization process.

In Section 2, we begin by defining concepts of local semi-simplicity and bounded local semi-simplicity of group representations. Theorem 1 states that bounded local semi-simplicity of a group representation is necessary and sufficient for the representation to be similar to a unitary representation. Several different forms of a conjecture concerning the group question are discussed, with the aid of Theorem 1. To study the operator algebra question, we develop a device for measuring the deviation of a set of vectors from being an orthonormal set. After stating a condition for a representation of a C^*-

algebra to be similar to a * representation in terms of the group condition (Theorem 1), we employ this device to give a more delicate criterion for topological semi-simplicity of an algebra of operators. In the concluding section, we discuss some extensions of the results stated, examples, and a class of natural questions an affirmative answer to any of which would yield the fact that all bounded operators on a Hilbert space have non-trivial, closed, invariant subspaces.

2. Conditions for topological semi-simplicity. The following definition contains a description of local behavior of a group representation, which is necessary and sufficient for the group representation to be similar to a unitary representation. The statement and proof of this fact are contained in Theorem 1.

DEFINITION 1. *A representation* $g \rightarrow A_g$ *of the group* G *by bounded operators* A_g *on the Banach space* $\mathfrak{B}$ *is said to be "locally semi-simple" when, for each finite set* $x_1, \cdots, x_n$ *of vectors in* $\mathfrak{B}$ *and* $g_1, \cdots, g_n$ *of elements in* G, *one can find a linear transformation* S *defined on the finite-dimensional vector space* $\mathcal{V}$ *generated by* $x_1, \cdots, x_n$, $A_{g_1}x_1, \cdots, A_{g_n}x_n$ *such that* $\| Sx_i \| = \| SA_{g_i}x_i \|$; $i = 1, \cdots, n$. *If there exists a constant* M *such that* S *can always be chosen satisfying*

$$1/M \leqq \inf \{ \| Sx \| : x \in \mathcal{V}, \| x \| = 1 \};$$
$$M \geqq \sup \{ \| Sx \| : x \in \mathcal{V}, \| x \| = 1 \},$$

we say that the representation is "boundedly locally semi-simple."

THEOREM 1. *A representation* $g \rightarrow A_g$ *of the group* G *by bounded operators* A_g *on a Hilbert space* $\mathfrak{H}$ *is boundedly locally semi-simple if and only if it is similar to a unitary representation.*

Proof. The necessity of the condition is quite easy. Indeed, suppose that S is a bounded invertible operator on $\mathfrak{H}$ such that SA_gS^{-1} is unitary for each g in G. Let $x_1, \cdots, x_n$ in $\mathfrak{H}$ and $g_1, \cdots, g_n$ in G be given. Since SA_gS^{-1} is unitary we have

$$\| SA_{g_i}x_i \| = \| SA_{g_i}S^{-1}Sx_i \| = \| Sx_i \|; \qquad i = 1, \cdots, n,$$

which establishes the bounded local semi-simplicity of the given representation.

Suppose now that the given group representation is known to be boundedly locally semi-simple and that M is a bounding constant. We establish, in the succeeding lemmas, the existence of an invertible operator

P with $\|P\|$, $\|P^{-1}\|$ not exceeding M and such that $P^{-1}A_gP$ is unitary for each g in G.

LEMMA 2. *If* $\mathcal{V}$ *is a finite-dimensional subspace of* $\mathfrak{H}$ *there exists a Hilbert space norm* $\|\ \|'$ *on* $\mathcal{V}$ *such that* $\|A_gx\|' = \|A_{g'}x\|'$ *whenever* A_gx *and* $A_{g'}x$ *are in* $\mathcal{V}$ *and such that* $1/M\|y\| \leqq \|y\|' \leqq M\|y\|$ *for each vector* y *in* $\mathcal{V}$.

Proof. We endow the conjugate tensor product $\mathcal{V}\otimes\mathcal{V}$ (i.e., the tensor product which is conjugate linear in the second variable, $x\otimes \alpha y = \bar{\alpha}(x\otimes y)$) with the natural inner product derived from the inner product of $\mathfrak{H}$, i.e., we define $\langle x\otimes y, z\otimes w\rangle = (x,z)(w,y)$ and extend the domain of definition of this inner product to all of $\mathcal{V}\otimes\mathcal{V}$ by bilinearity. It is well-known that this process gives rise to an inner product (independent) of the representation of the elements involved as a sum of elements of the form $x\otimes y$—(see [3], for example).

Let $\mathcal{V}\otimes$ be the subspace of $\mathcal{V}\otimes\mathcal{V}$ generated by tensors of the form $x\otimes x$, and let $\mathcal{E}\otimes$ be the subspace of $\mathcal{V}\otimes$ generated by vectors of the form $A_gy\otimes A_gy - y\otimes y$, where y and A_gy are in $\mathcal{V}$. Choose a basis

$$A_{g_1}y_1\otimes A_{g_1}y_1 - y_1\otimes y_1, \cdots, A_{g_m}y_m\otimes A_{g_m}y_m - y_m\otimes y_m$$

for $\mathcal{E}\otimes$. By the bounded local semi-simplicity of the representation $g\to A_g$, we can find a linear transformation S such that $\|SA_{g_i}y_i\| = \|Sy_i\|$; $i=1,\cdots,m$ and

$$1/M \leqq \inf\{\|Sy\| : y \text{ in } \mathcal{V}', \|y\|=1\};$$
$$M \geqq \sup\{\|Sy\| : y \text{ in } \mathcal{V}', \|y\|=1\},$$

where $\mathcal{V}'$ is the subspace of $\mathcal{V}$ generated by $y_1,\cdots,y_m$, $A_{g_1}y_1,\cdots,A_{g_m}y_m$. By defining S to be a suitable scalar (between $1/M$ and M) on $\mathcal{V}\ominus\mathcal{V}'$, we obtain a linear transformation, which we denote again by S, defined on all of $\mathcal{V}$ and satisfying the same conditions as above with $\mathcal{V}$ replacing $\mathcal{V}'$ (no difficulties can arise if we choose S so that $S(\mathcal{V}')=\mathcal{V}'$ by composing the original S with a unitary transformation).

Let $\tilde{x} = \sum_{j=1}^{n} S^*x_j\otimes S^*x_j$, where $x_1,\cdots,x_n$ is an orthonormal basis for $S(\mathcal{V})$. Observe that

$$\langle\tilde{x}, A_{g_i}y_i\otimes A_{g_i}y_i - y_i\otimes y_i\rangle = \sum_{j=1}^{n}(S^*x_j, A_{g_i}y_i)(A_{g_i}y_i, S^*x_j)$$
$$-\sum_{j=1}^{n}(S^*x_j, y_i)(y_i, S^*x_j) = \sum_{j=1}^{n}|(x_j, SA_{g_i}y_i)|^2 - \sum_{j=1}^{n}|(x_j, Sy_i)|^2$$
$$= \|SA_{g_i}y_i\|^2 - \|Sy_i\|^2 = 0,$$

as follows from the fact that $SA_{g_i}y_i$, Sy_i are in $S(\mathcal{V})$, the Parseval equality, and the choice of S. Consequently $\tilde{x}$ is orthogonal to $\mathcal{E}\otimes$. We define a new inner product on $\mathcal{V}$ by means of $(x, y)' = \langle x \otimes y, \tilde{x} \rangle$ so that the new Hilbert norm on $\mathcal{V}$ satisfies

$$\| y \|' = \langle y \otimes y, \tilde{x} \rangle^{\frac{1}{2}} = [\sum_{j=1}^{n} (S^* x_j, y)(y, S^* x_j)]^{\frac{1}{2}}$$

$$= [\sum_{j=1}^{n} |(x_j, Sy)|^2]^{\frac{1}{2}} = \| Sy \|.$$

It follows at once from this last equality that $1/M \| y \| \leqq \| y \|' \leqq M \| y \|$. If now y and $A_g y$ are in $\mathcal{V}$ then $A_g y \otimes A_g y - y \otimes y$ is in $\mathcal{E}\otimes$ so that

$$0 = (A_g y \otimes A_g y - y \otimes y, \tilde{x}) = (A_g y \otimes A_g y, \tilde{x}) - (y \otimes y, \tilde{x})$$

$$= (\| A_g y \|')^2 - (\| y \|')^2$$

or $\| y \|' = \| A_g y \|'$. If $A_g x$ and $A_{g'} x = A_{g'} A_{g^{-1}} (A_g x)$ are in $\mathcal{V}$, then $\| A_g x \|' = \| A_{g'} x \|'$, which completes the proof of this lemma.

The following lemma allows us to pass from our finite-dimensional information to information about the full space on which the A_g operate.

LEMMA 3. *If $\mathfrak{B}$ is a Banach space and ρ is a partition function on $\mathfrak{B}$ such that on each finite-dimensional subspace $\mathfrak{B}_1$ of $\mathfrak{B}$ one can introduce a norm $\| \ \|'$ in which $\mathfrak{B}_1$ is a Hilbert space, each partition class intersected with $\mathfrak{B}_1$ lies on the shell of some sphere center at 0 in the norm $\| \ \|'$, and there exists a constant M (depending upon ρ) such that*

$$1/M \| x \| \leqq \| x \|' \leqq M \| x \|$$

for each x in $\mathfrak{B}_1$ (where $\| \ \|$ is the norm on $\mathfrak{B}$); then the underlying vector space of $\mathfrak{B}$ admits a norm, equivalent to the original norm, in which it is a Hilbert space and such that the partition classes of ρ each lie on the shell of some sphere center at 0 relative to the new norm.

Proof. We form a product of intervals with $\mathfrak{B}$ as the indexing family. To each point x in $\mathfrak{B}$, we make correspond the closed interval $[\| x \|/M, M \| x \|]$ (thus to 0 in $\mathfrak{B}$ we make correspond the number 0). Denote by X the Cartesian product

$$\prod_{x \in \mathfrak{B}} I_x = \prod_{x \in \mathfrak{B}} [\| x \|/M, M \| x \|].$$

We consider X in its standard product topology, in which it is compact, where each I_x is given its usual metric topology. Let $\mathfrak{B}_1$ be a finite-dimensional subspace of $\mathcal{B}$, and let $X(\mathcal{B}_1)$ be the set of points of X which as functions restricted to $\mathcal{B}_1$ give rise to a Hilbert space norm on $\mathcal{B}_1$ which is constant on the partition classes of ρ intersected with $\mathcal{B}_1$. We shall show presently that $X(\mathcal{B}_1)$ is a closed subset of X. Assume, for the moment, that we have proved this fact. The sets $X(\mathcal{B}_1)$ have the finite intersection property ($\mathcal{B}_1$ ranging over the finite-dimensional subspaces of $\mathcal{B}$). Indeed, let $\mathcal{B}_1, \cdots, \mathcal{B}_n$ be a (finite) set of finite-dimensional subspaces of $\mathcal{B}$ and let $\mathcal{B}_0$ be the (finite-dimensional) subspace they generate. By assumption, we can find a Hilbert space norm $\| \ \|_0$ on $\mathcal{B}_0$ which is constant on the partition classes of ρ intersected with $\mathcal{B}_0$, and which satisfies the inequality

$$1/M \| x \| \leqq \| x \|_0 \leqq M \| x \|$$

for each x in $\mathcal{B}_0$. The function which assigns to each x not in $\mathcal{B}_0$ the value $M \| x \|$ and to each x in $\mathcal{B}_0$ the value $\| x \|_0$ lies in $X(\mathcal{B}_0)$ which is clearly contained in $\bigcap_{i=1}^{n} X(\mathcal{B}_i)$. It now follows from the compactness of X (and our assumption that the sets $X(\mathcal{B}_1)$ are closed in X) that the intersection of all the sets $X(\mathcal{B}_1)$ is not empty. Let $\| \ \|'$ be a function on $\mathcal{B}$ in this intersection. Then, on each finite-dimensional subspace of $\mathcal{B}$, $\| \ \|'$ induces a Hilbert space norm. It is immediate that $\| \ \|'$ satisfies the norm axioms and the Parallelogram Law on $\mathcal{B}$ as well as being constant on the partition classes of ρ, so that $\| \ \|'$ is our desired Hilbert space norm on $\mathcal{B}$. Of course $1/M \| x \| \leqq \| x \|' \leqq M \| x \|$, since $\| \ \|'$ is in X. It remains to prove that the sets $X(\mathcal{B}_1)$ are closed in X. We shall omit this proof, however, since it is a standard approximation argument of the type employed in the proof of the w^*-compactness of the unit sphere in the conjugate space of a normed linear space.

Proof of Theorem 1. As partition function ρ on $\mathcal{H}$ we take the map which assigns to each vector x in $\mathcal{H}$ the set of vectors $\{A_g x : g \text{ in } G\}$. Since the family of operators $\{A_g\}$ forms a group, this map defines a partition function on $\mathcal{H}$. Lemma 2 establishes the hypothesis of Lemma 3 with this partition function and $\mathcal{H}$ for $\mathcal{B}$, so that we can conclude the existence of a norm $\| \ \|'$ on $\mathcal{H}$ in which $\mathcal{H}$ is a Hilbert space and such that $\| A_{g_1} x \|' = \| A_{g_2} x \|'$ for each x in $\mathcal{H}$ and g_1, g_2 in G. In particular $\| x \|' = \| A_g x \|'$ so that each operator A_g is isometric with respect to the norm $\| \ \|'$. Moreover, $\| \ \|'$ can be so chosen that $\| x \|/M \leqq \| x \|' \leqq M \| x \|$ for each vector x in $\mathcal{H}$.

Let $x_1, \cdots$ be an orthonormal basis for $\mathcal{H}$ with respect to the norm $\| \ \|$

(and associated inner product $(\,,\,)$), and let $y_1, \cdots$ be an orthonormal basis for $\mathfrak{H}$ with respect to the norm $\| \;\|'$ (and associated inner product $(\,,\,)'$). Define a linear transformation P of $\mathfrak{H}$ into itself by $Px_i = y_i$; $i = 1, \cdots$. Then

$$(x, y) = (\sum_i (x, x_i) x_i, \sum_i (y, x_i) x_i) = \sum_i (x, x_i)(x_i, y)$$
$$= (\sum_i (x, x_i) y_i, \sum_i (y, x_i) y_i)' = (Px, Py)'.$$

Of course $(P^{-1}x, P^{-1}y) = (x, y)'$, substituting $P^{-1}x$ for x and $P^{-1}y$ for y throughout the above equality. We assert that $P^{-1}A_gP$ is a unitary operator on $\mathfrak{H}$ with respect to the norm $\| \;\|$ for each g in G. Indeed,

$$(P^{-1}A_gPx, P^{-1}A_gPy) = (A_gPx, A_gPy)' = (Px, Py)' = (x, y).$$

We note in conclusion that $\| P \|$, $\| P^{-1} \|$ do not exceed M. In fact, if $x = \sum_i \alpha_i x_i$ with $1 = \| x \|^2 = \sum_i | \alpha_i |^2$ is given, then $\| Px \|' = \| \sum_i \alpha_i y_i \|' = 1$ and $\| Px \| / M \leqq \| Px \|' \leqq M \| Px \|$, so that $1/M \leqq \| Px \| \leqq M$.

There are several ways of formulating a conjecture concerning the classical question of whether or not each bounded representation of a group is similar to a unitary representation.

CONJECTURE A. *Every bounded representation of a group by operators on a Hilbert space is similar to a unitary representation.*

CONJECTURE B. *There exists a function f from the positive reals to the positive reals with the property that for each bounded representation $g \to A_g$ with bound M, of a group G by operators on a Hilbert space one can find an invertible operator P such that $P^{-1}A_gP$ is unitary for each g in G and such that $\| P \|$, $\| P^{-1} \|$ do not exceed $f(M)$.*

CONJECTURE C. *Same as* B *with f as the identity transform.*

Each of the above conjectures is clearly stronger than the preceding one. We shall show that B is actually equivalent to A in the next lemma.

LEMMA 4. *Conjecture* A *is equivalent to Conjecture* B.

Proof. Clearly B implies A. Suppose now that A is true. If B is false there exists a sequence of groups $G_1, G_2, \cdots$ and a sequence of representations $g^{(1)} \to A_{g^{(1)}}$, $g^{(2)} \to A_{g^{(2)}}, \cdots$ of these groups, respectively, each with bound M and such that if $N_i = \inf\{\max(\| P_i \|, \| P_i^{-1} \|) : P_i^{-1} A_{g^{(i)}} P_i$ unitary for each $g^{(i)}$ in $G_i\}$ then $\lim_i N_i = \infty$. Let $G = G_1 \otimes G_2 \otimes \cdots$ be the weak direct sum of the groups $G_1, G_2, \cdots$, and let $g \rightarrow A_g$ be the direct sum of the representations $g^{(1)} \rightarrow A_{g^{(1)}}, \cdots$. The representation $g \rightarrow A_g$ of G has bound M. Assuming A, we can find an operator P such that $P^{-1}A_gP$ is unitary for each g in G. Restricted to each direct summand, this similarity induces similarities of all the representations $g^{(i)} \rightarrow A_{g^{(i)}}$, each similarity with bound not greater than $\max(\| P \|, \| P^{-1} \|)$—a contradiction. Hence A implies B.

THEOREM 5. *If* B *is true for the free groups on finitely many generators then* B *is true for all groups.*

Proof. Let $g \rightarrow A_g$ be a representation of G with bound M. We shall show that this representation is boundedly locally semi-simple with bounding constant $f(M)$. In fact, let $x_1, \cdots, x_n$ in $\mathfrak{H}$ and $g_1, \cdots, g_n$ in G be given. The group G_n generated by $g_1, \cdots, g_n$ is the homomorphic image of F_n, the free group on n generators. Thus the representation $g \rightarrow A_g$ of G restricted to G_n gives rise to a representation of F_n with bound M which, by hypothesis, is similar to a unitary representation via an operator P with $\| P \|$, $\| P^{-1} \|$ not exceeding $f(M)$. As in the proof of Theorem 1, we now conclude that $\| PA_{g_i}x_i \| = \| Px_i \|$, $i = 1, \cdots, n$; so that the representation is boundedly locally semi-simple with bounding constant $f(M)$. Hence, by Theorem 1, the representation $g \rightarrow A_g$ is similar to a unitary representation via a T such that $\| T \|$, $\| T^{-1} \|$ do not exceed $f(M)$. Thus B follows for all groups.

Note that the proof of Lemma 4 shows that assuming A for the class of groups generated by no more than a countable number of elements implies B for this class (since the group G constructed in the proof would be in this class). Now every group in this class is the homomorphic image of the free group on countably many generators, F_∞, so that assuming A for F_∞ implies A for all groups with a countable number of generators and hence B for the free groups on finitely many generators. With the theorem just proved, this yields:

COROLLARY 6. *If* A *holds for the free group on a countable infinity of generators then* A *and hence* B *holds for all groups.*

We turn our attention now to the question of topological semi-simplicity of algebras of operators. In Theorem 8, we state a necessary and sufficient condition for a representation of a C^*-algebra to be similar to a * representation. Before stating this result, however, it is necessary to introduce some geometrical concepts. In particular we must associate to each configuration of vectors an object which measures its deviation from being an orthonormal set. To this end, we introduce an "inner product" between two sets of n vectors in $\mathcal{H}$. This inner product has as its range of values, operators on $\mathcal{H}$.

DEFINITION 2. *If* $\tilde{x} = (x_1, \cdots, x_n)$, $\tilde{y} = (y_1, \cdots, y_n)$ *are two n-tuples of vectors in* $\mathcal{H}$ *with* $\mathcal{V}$, $\mathcal{W}$ *the spaces generated by* $\{x_1, \cdots, x_n\}$; $\{y_1, \cdots, y_n\}$, *respectively, we denote by* $\langle \tilde{x}, \tilde{y} \rangle$ *the operator on* $\mathcal{H}$ *defined as follows. Let* C^n *be the space of n-tuples of complex numbers with the usual inner product and let* $e_1, \cdots, e_n$ *be the basis* $(1, 0, \cdots, 0), \cdots, (0, \cdots, 0, 1)$. *Let* P *be the map of* C^n *into* $\mathcal{V}$ *determined by* $P(e_i) = x_i$, $i = 1, \cdots, n$, *and let* Q *be the map of* C^n *into* $\mathcal{W}$ *determined by* $Q(e_i) = y_i$. *By* Q^* *we mean the adjoint map to* Q, *from* $\mathcal{H}$ *into* C^n *(characterized by* $(Q^*x, a) = (x, Qa)$, *where* x *is an arbitrary vector in* $\mathcal{H}$, a *in* C^n, *the first inner product is taken in* C^n, *and the second in* $\mathcal{H}$). *Then* $\langle \tilde{x}, \tilde{y} \rangle = PQ^*$.

We note some of the properties of $\langle\ ,\ \rangle$. As a function on the product of $\mathcal{H}_n$ (the n-fold direct sum of $\mathcal{H}$ with itself) with $\mathcal{H}_n$, this inner product is conjugate bilinear. Indeed with $\tilde{x} = (x_1, \cdots, x_n)$, $\tilde{x}' = (x_1', \cdots, x_n')$, $\tilde{y} = (y_1, \cdots, y_n)$, $\tilde{y}' = (y_1', \cdots, y_n')$, and $\langle \tilde{x}, \tilde{y} \rangle = PQ^*$, $\langle \tilde{x}', \tilde{y} \rangle = P'Q^*$, $\langle \tilde{x}, \tilde{y}' \rangle = PQ'^*$ we have

$$\langle \tilde{x} + \tilde{x}', \alpha\tilde{y} \rangle = (P + P')(\alpha Q)^* = \bar{\alpha}(PQ^* + P'Q^*) = \bar{\alpha}(\langle \tilde{x}, \tilde{y} \rangle + \langle \tilde{x}', \tilde{y} \rangle)$$

and similarly

$$\langle \alpha\tilde{x}, \tilde{y} + \tilde{y}' \rangle = \alpha(\langle \tilde{x}, \tilde{y} \rangle + \langle \tilde{x}, \tilde{y}' \rangle), \qquad \langle \tilde{x}, \tilde{y} \rangle = PQ^* = (QP^*)^* = \langle \tilde{y}, \tilde{x} \rangle^*.$$

With the notation as in the definition, we see that the range of $\langle \tilde{x}, \tilde{y} \rangle$ is contained $\mathcal{V}$ and that the range of $\langle \tilde{x}, \tilde{y} \rangle^* = \langle \tilde{y}, \tilde{x} \rangle = QP^*$ is contained in $\mathcal{W}$. Since the null space of an operator is the complement of the range of its adjoint, we have that the complement of $\mathcal{W}$ in $\mathcal{H}$ is the null space of $\langle \tilde{x}, \tilde{y} \rangle$. Thus, effectively, $\langle \tilde{x}, \tilde{y} \rangle$ is a transformation from $\mathcal{W}$ to $\mathcal{V}$.

We compute the transformation $\langle \tilde{x}, \tilde{y} \rangle$ precisely. With the notation above we have, for z in $\mathcal{H}$:

$$(Q^*z, e_i) = (z, Qe_i) = (z, y_i),$$

so that $Q^*z = \sum_{i=1}^{n} (z, y_i) e_i$, hence $PQ^*z = \sum_{i=1}^{n} (z, y_i) x_i$. Thus $\langle \tilde{x}, \tilde{y} \rangle$ can be expressed symbolically as $(\ , y_1)x_1 + \cdots + (\ , y_n)x_n$. It is immediate from this, that $\langle \tilde{x}, \tilde{x} \rangle$ is a positive operator on $\mathcal{H}$ (as it is from the expression PP^* for $\langle \tilde{x}, \tilde{x} \rangle$), and as such has a (unique) positive square root. We denote this square root by $\langle \tilde{x} \rangle$ and refer to it as "the geometrical norm of $\tilde{x}$ (or of the configuration $x_1, \cdots, x_n$)." The fact that $\langle \tilde{x} \rangle$ is the identity operator on the n-dimensional space $\mathcal{V}$ (so that $x_1, \cdots, x_n$ are linearly independent, in particular) is equivalent to

$$\langle\tilde{x}\rangle^2 = \langle\tilde{x}, \tilde{x}\rangle = (\ , x_1)x_1 + \cdots + (\ , x_n)x_n$$

being the identity transformation on $\mathcal{V}$, which is equivalent to $x_1, \cdots, x_n$ being an orthonormal frame. The spread of the spectrum of $\langle\tilde{x}\rangle$, in general, is a measure of how much $x_1, \cdots, x_n$ deviates from being a scalar multiple of an orthonormal set. In the one-dimensional case, i. e., with x, y vectors in $\mathcal{H}$, we have $\langle x, y\rangle = (\ , y)x$. If we restrict this operator to the one-dimensional space generated by x, it becomes multiplication by (x, y), the usual inner product of x and y.

Suppose, now, that $\mathfrak{A}$ is a C^*-algebra and ϕ is a representation (not necessarily $*$ preserving) of $\mathfrak{A}$ by operators on a Hilbert space $\mathcal{H}$. Employing Theorem 1, we obtain the following criterion for ϕ to be similar to a $*$ representation.

THEOREM 7. *If $\mathcal{U}$ is the unitary group of the C^*-algebra $\mathfrak{A}$, a necessary and sufficient condition for a representation ϕ of $\mathfrak{A}$ by operators on a Hilbert space $\mathcal{H}$ to be similar to a $*$ representation of $\mathfrak{A}$ is that ϕ restricted to $\mathcal{U}$ be a boundedly, locally semi-simple group representation of $\mathcal{U}$.*

Proof. If ϕ is similar to a $*$ representation there exists an operator P on $\mathcal{H}$ such that $P^{-1}\phi(U)P$ is unitary, for each U in $\mathcal{U}$. Thus ϕ restricted to $\mathcal{U}$ is similar to a unitary representation of $\mathcal{U}$; and ϕ is boundedly, locally semi-simple, by Theorem 1. On the other hand, if ϕ is boundedly, locally semi-simple as a representation of $\mathcal{U}$ then, by Theorem 1, there exists an operator P on $\mathcal{H}$ such that $P^{-1}(U)P$ is unitary for each unitary operator U in $\mathfrak{A}$. It now follows that $A \rightarrow P^{-1}\phi(A)P$ is a $*$ representation of $\mathfrak{A}$. Indeed, the given map is an algebraic isomorphism of $\mathfrak{A}$. Suppose A is a self-adjoint operator in $\mathfrak{A}$ of norm not exceeding 1. Then $A = \frac{1}{2}(U_1 + U_2)$ where $U_1 = A + i(I - A^2)^{\frac{1}{2}}$ and $U_2 = A - i(I - A^2)^{\frac{1}{2}}$ are unitary operators in $\mathfrak{A}$. Thus $P^{-1}\phi(A)P = \frac{1}{2}[P^{-1}\phi(U_1)P + P^{-1}\phi(U_2)P]$ with $P^{-1}\phi(U_1)P$ and $P^{-1}\phi(U_2)P$ unitary, so that

$$[P^{-1}\phi(U_i)P]^* = [P^{-1}\phi(U_i)P]^{-1} = P^{-1}\phi(U_i^{-1})P = P^{-1}\phi(U_i^*)P; \quad i = 1, 2.$$

Thus

$$[P^{-1}\phi(A)P]^* = P^{-1}\phi(\tfrac{1}{2}(U_1^* + U_2^*))P = P^{-1}\phi(A^*)P = P^{-1}\phi(A)P,$$

so that $A \rightarrow P^{-1}\phi(A)P$ takes self-adjoint operators in $\mathfrak{A}$ into self-adjoint operators, and, therefore, is a $*$ representation of $\mathfrak{A}$.

Making use of the foregoing concept of inner product between sets of

n vectors, it is possible to give a more delicate analysis to the question of which representations of C^*-algebras are similar to * representations. If ϕ is a representation of a C^*-algebra $\mathfrak{A}$ by operators on a Hilbert space $\mathcal{H}'$ ($\mathfrak{A}$ acts on $\mathcal{H}$) and $\phi(A)x' = 0$ for some unit vector x' in $\mathcal{H}'$, then for each positive ϵ one can find a unit vector x in $\mathcal{H}$ such that $\| Ax \| < \epsilon$. Indeed, if $\phi(A_i)x' = 0$, $i = 1, \cdots, n$ one can choose the unit vector x so that $\| A_i x \| < \epsilon$, $i = 1, \cdots, n$, i.e., the relations $\phi(A_i)x' = 0$ can be "approximately duplicated" with $\mathfrak{A}$ and $\mathcal{H}$ via ϕ. In fact, the set of operators A such that $\phi(A)x' = 0$ forms a proper left ideal $\mathfrak{J}$ in $\mathfrak{A}$ (proper, since $\phi(I) = I$). If for each unit vector x in $\mathcal{H}$ one of $A_i x$ has norm not less than ϵ then $T = \sum A_i^* A_i \geqq \epsilon I$. But T is then invertible and in $\mathfrak{J}$. This contradiction implies the existence of the desired unit vector x. Given $\epsilon > 0$, vectors $x_1', \cdots, x_n'$ in $\mathcal{H}'$ such that $\sum \| x_i' \|^2 = 1$, and relations $\sum_{i=1}^{n} \phi(A_{hi})x_i' = 0$, $h = 1, \cdots, m$; it is even possible to find vectors $x_1, \cdots, x_n$ in $\mathcal{H}$ with $\sum \| x_i \|^2 = 1$ such that $\| \sum_{i=1}^{n} A_{hi} x_i \| < \epsilon$, $h = 1, \cdots, m$. This can be done by working with the $n \times n$ matrix algebras over $\mathfrak{A}$ and $\phi(\mathfrak{A})$ as we did above with $\mathfrak{A}$ and $\phi(\mathfrak{A})$ themselves. On the other hand, suppose the relations $\sum_{i=1}^{n} \phi(A_{hi})x_i' = 0$, $h = 1, \cdots, m$, subsist with $x_1', \cdots, x_n'$ an orthonormal set in $\mathcal{H}'$; is it possible to choose $x_1, \cdots, x_n$ an orthonormal set in $\mathcal{H}$ such that $\| \sum_{i=1}^{n} A_{hi} x_i \| < \epsilon$, $h = 1, \cdots, m$? This is not necessarily possible on two grounds; a multiplicity consideration, or more simply, the dimension of $\mathcal{H}$ may not be large enough to accommodate an orthonormal set with n vectors, secondly, it is too much to ask for orthonormality of $x_1, \cdots, x_n$ in light of the fact that ϕ may not be a * representation (Theorem 8 shows that if it is possible to choose $x_1, \cdots, x_n$ an orthonormal set then ϕ is already a * representation). The multiplicity question can be avoided by asking whether or not a * representation ψ of $\mathfrak{A}$ can be found (once the relations $\sum_{i=1}^{n} \phi(A_{hi})x_i' = 0$ and $\epsilon > 0$ are given) such that $\| \sum_{i=1}^{n} \psi(A_{hi}) x_i \| < \epsilon$. As for the orthonormality question, can we at least find bounds, dependent upon the representation ϕ alone, for the distortion of $x_1, \cdots, x_n$ from being an orthonormal set? The technique for measuring this distortion has just been developed. It is not difficult to see that if ϕ is similar to a * representation ψ via an operator P, then ψ will serve for the exact duplication of all relations with the distortion bounded by $\max(\| P \|, \| P^{-1} \|)$ (this will be done in detail in the necessity portion of Theorem 8). These considerations lead us to:

DEFINITION 3. *Let* ϕ *be a representation of the* C^**-algebra* $\mathfrak{A}$ *by operators on the Hilbert space* $\mathcal{H}$, *and let* $\tilde{x} = (x_1, \cdots, x_n)$ *be an n-tuple consisting of vectors* $x_1, \cdots, x_n$ *which form an orthonormal set in* $\mathcal{H}$ *such that* $(\phi(A_{ij}))\tilde{x} = 0$, *where* $(\phi(A_{ij}))$ *is an* $n \times n$ *matrix whose entries are operators in* $\phi(\mathfrak{A})$. *Let* ψ *be a* $*$ *representation of* $\mathfrak{A}$ *by operators on a Hilbert space* $\mathcal{H}'$, *and let* $\tilde{x}' = (x_1', \cdots, x_n')$ *be an n-tuple of vectors in* $\mathcal{H}'$ *such that* $\| (\psi(A_{ij})\tilde{x}' \| < \epsilon$, *where* ϵ *is some positive number and the spectrum of* $\langle \tilde{x}' \rangle$, *as an operator on the space generated by* $x_1', \cdots, x'_n$, *is contained in the interval* $[k, K]$. *We say, then, that "*$\| (\psi(A_{ij}))\tilde{x}' \| < \epsilon$ *is a self-adjoint* ϵ *cover of the relations* $(\phi(A_{ij}))\tilde{x} = 0$ *with distortion in* $[k, K]$*." If there exist constants* k, K, $(K > k > 0)$ *such that each relation of the form* $(\phi(A_{ij}))\tilde{x} = 0$, *with* $\tilde{x}$ *as above and* (A_{ij}) *a positive operator (in the* C^**-algebra consisting of* $n \times n$ *matrices over* $\mathfrak{A}$), *has, for each positive* ϵ, *a self-adjoint* ϵ *cover with distortion in* $[k, K]$, *we say that "the representation* ϕ *has a self-adjoint cover (with distortion in* $[k, K]$)*."*

We have not made the definition of a representation having a self-adjoint cover as restrictive as we might, in that we require only relations coming from positive $n \times n$ matrices to have self-adjoint ϵ covers. This is all that is needed for each relation to have a cover. It might seem more natural to use the phrase "ϕ has a self-adjoint cover" to mean that for each ϵ there is a self-adjoint representation which serves as a self-adjoint ϵ cover of ϕ for all relations. That this actually follows from the weaker condition used and, indeed, that there is a self-adjoint representation which works for all positive ϵ and all relations is the substance of:

THEOREM 8. *A necessary and sufficient condition for a representation* ϕ *of a* C^**-algebra* $\mathfrak{A}$ *by operators on a Hilbert space* $\mathcal{H}$ *to be similar to a* $*$ *representation is that* ϕ *have a self-adjoint cover. If the distortion is in* $[k, K]$ *then a similarity can be effected by a positive operator with spectrum in* $[k, K]$.

Proof. The necessity presents little difficulty. Suppose that there exists an invertible operator T on $\mathcal{H}$ such that $A \to T\phi(A)T^{-1}$ is a $*$ representation of $\mathfrak{A}$, and let $M = \max(\| T \|, \| T^{-1} \|)$. If $(\phi(A_{ij}))\tilde{x} = 0$ is some relation, with $\tilde{x} = (x_1, \cdots, x_n)$, $x_1, \cdots, x_n$ an orthonormal set, then $(T\phi(A_{ij})T^{-1})\tilde{x}' = 0$ is a self-adjoint ϵ cover for this relation (all $\epsilon > 0$), where $\tilde{x}' = (Tx_1, \cdots, Tx_n)$, and where the distortion lies in $[1/M, M]$. Indeed, that $(T\phi(A_{ij})T^{-1})\tilde{x}' = 0$ with the given $\tilde{x}'$ is immediate. Let P be the linear transformation from C^n

into $\mathcal{H}$ defined by $Pe_i = x_i$ (see Definition 2), and let E be the projection on the space generated by $x_1, \cdots, x_n$. Then

$$\langle \tilde{x}', \tilde{x}' \rangle = TPP^*T^* = TET^* = (TE)(TE)^*.$$

Thus $\| \langle \tilde{x}', \tilde{x}' \rangle \| = \| TE \|^2 \leqq \| T \|^2 \leqq M^2$. Moreover

$$\begin{aligned} &\inf\{(\langle \tilde{x}', \tilde{x}' \rangle x, x) : \| x \| = 1, x \text{ in } TE\mathcal{H}\} = \inf\{\| (TE)^* x \|^2\} \\ &\quad = \inf\{\| ET^*Ty \|^2 : y \text{ in } E\mathcal{H}, \| Ty \| = 1\} \geqq \inf\{(T^*Ty, y/\| y \|)^2\} \\ &\quad = \inf\{\| Ty \|^4/\| y \|^2\} = \inf\{1/\| y \|^2 : y \text{ in } E\mathcal{H}, \| Ty \| = 1\} \geqq 1/M^2. \end{aligned}$$

Thus the spectrum of $\langle \tilde{x}', \tilde{x}' \rangle$ as an operator on $TE\mathcal{H}$ lies in $[1/M^2, M^2]$ so that the spectrum of $\langle \tilde{x}' \rangle$ lies in $[1/M, M]$, and $(T\phi(A_{ij})T^{-1})\tilde{x}' = 0$ is a self-adjoint ϵ cover (all $\epsilon > 0$) for $(\phi(A_{ij}))\tilde{x} = 0$ with distortion in $[1/M, M]$. In connection with foregoing inequalities, note that y is in $E\mathcal{H}$ so that the length of the projection of T^*Ty upon $E\mathcal{H}$ is not less than the length of the projection of T^*Ty upon the subspace generated by y (this length being $(T^*Ty, y/\| y \|)$).

Suppose now that the map ϕ has a self-adjoint cover. As in Theorem 1, we show that each finite-dimensional subspace $\mathcal{V}$ of $\mathcal{H}$ admits a Hilbert space norm $\| \ \|'$ such that $\| \phi(U)x \|' = \| \phi(V)x \|'$ when U, V are unitary operators in $\mathfrak{A}$ with $\phi(U)x, \phi(V)x$ in $\mathcal{V}$, and such that $k \| y \| \leqq \| y \|' \leqq K \| y \|$ for each y in $\mathcal{V}$. Following Theorem 1, form the conjugate tensor product $\mathcal{V} \otimes \mathcal{V}$ of $\mathcal{V}$ with itself and endow it with the unitary structure described in that theorem. Let $\mathcal{V} \otimes$ be the subspace of $\mathcal{V} \otimes \mathcal{V}$ generated by tensors of the form $x \otimes x$ and $\mathcal{E} \otimes$ the subspace generated by elements $\phi(U)x \otimes \phi(U)x - x \otimes x$, where U is a unitary operator in $\mathfrak{A}$ and $x, \phi(U)x$ are in $\mathcal{V}$. Choose a basis $\phi(U_1)y_1 \otimes \phi(U_1)y_1 - y_1 \otimes y_1, \cdots, \phi(U_m)y_m \otimes \phi(U_m)y_m - y_m \otimes y_m$ for $\mathcal{E} \otimes$ and an orthonormal basis $x_1, \cdots, x_n$ for $\mathcal{V}$. We have

$$\phi(U_i)y_i = \sum_{j=1}^{n} \beta'_{ij} x_j \text{ and } y_i = \sum_{j=1}^{n} \beta_{ij} x_j; \qquad i = 1, \cdots, m,$$

so that

$$0 = \sum_{j=1}^{n} (\beta_{ij}\phi(U_i) - \beta'_{ij}) x_j = \sum_{j=1}^{n} \phi(\beta_{ij}U_i - \beta'_{ij}I) x_j; \qquad i = 1, \cdots, m.$$

Let a positive integer r and a positive number δ be given. We wish to establish the existence of a * representation ψ of $\mathfrak{A}$ as operators on a Hilbert space $\mathcal{H}'$ and vectors $x_1', \cdots, x_n'$ in $\mathcal{H}'$ such that $\langle \tilde{x}' \rangle$ has its spectrum in $[k, K]$, with $\tilde{x}' = (x_1', \cdots, x_n')$, and such that

$$\| \sum_{j=1}^{n} \psi(\beta_{ij}U_i - \beta'_{ij}I) x_j' \| < \delta; \qquad i = 1, \cdots, m.$$

We write A_{hj} for $\beta_{hj}U_h - \beta'_{hj}I$; $h = 1, \cdots, m$ and $\phi(A)^{\sim}$ for the $n \times n$ matrix whose i, j entry is the operator $\sum_{h=1}^{m} \phi(A_{hi}{}^*)\phi(A_{hj})$. Now $\phi(A)^{\sim}\tilde{x} = 0$, where $\tilde{x} = (x_1, \cdots, x_n)$, for

$$\phi(A)^{\sim} = \sum_{h=1}^{m} \begin{bmatrix} \phi(A_{h1}{}^*) & 0 \cdots 0 \\ \vdots & \vdots \\ \phi(A_{hn}{}^*) & 0 \cdots 0 \end{bmatrix} \cdot \begin{bmatrix} \phi(A_{h1}) \cdots \phi(A_{hn}) \\ 0 \quad \cdots \quad 0 \\ \vdots \qquad\qquad \vdots \\ 0 \quad \cdots \quad 0 \end{bmatrix}$$

and

$$\begin{bmatrix} \phi(A_{h1}) \cdots \phi(A_{hn}) \\ 0 \quad \cdots \quad 0 \\ \vdots \qquad\qquad \vdots \\ 0 \quad \cdots \quad 0 \end{bmatrix} \tilde{x} = 0,$$

$h = 1, \cdots, m$. By hypothesis on ϕ, the relation $\phi(A)^{\sim}\tilde{x} = 0$ has a self-adjoint δ cover with distortion in $[k, K]$. Let ψ be a representation of $\mathfrak{A}$ by operators on the Hilbert space $\mathcal{H}'$ and let $\tilde{x}' = (x_1', \cdots, x_n')$ be a vector such that $\langle \tilde{x}' \rangle$ has spectrum in $[k, K]$ and $\| \psi(A)^{\sim}\tilde{x}' \| < \delta^2/n^{\frac{1}{2}}K$ where $\psi(A)^{\sim}$ is the $n \times n$ matrix whose i, j entry is $\sum_{h=1}^{m} \psi(A_{hi}{}^*)\psi(A_{hi})$. In particular then,

$$(\psi(A)^{\sim}\tilde{x}', \tilde{x}') \leqq \| \psi(A)^{\sim}\tilde{x}' \| \cdot \| \tilde{x}' \| < \delta^2$$

for

$$\| \tilde{x}' \| = (\sum_{i=1}^{n} \| x_i' \|^2)^{\frac{1}{2}} \leqq (\sum_{i=1}^{n} K^2)^{\frac{1}{2}} = n^{\frac{1}{2}}K.$$

In fact, since $\langle \tilde{x}' \rangle$ has spectrum in $[k, K]$, $\langle \tilde{x}', \tilde{x}' \rangle = \langle \tilde{x}' \rangle^2$ has spectrum in $[k^2, K^2]$, so that

$$\| \langle \tilde{x}', \tilde{x}' \rangle \| = \| PP^* \| = \| P \|^2 \leqq K^2 \text{ and } \| x_i' \| = \| Pe_i \| \leqq K$$

(notation as in Definition 2). Now $(\psi(A)^{\sim}\tilde{x}', \tilde{x}') =$

$$\sum_{h=1}^{m} (\begin{bmatrix} \psi(A_{h1}{}^*) & 0 \cdots 0 \\ \vdots & \vdots \\ \psi(A_{hn}{}^*) & 0 \cdots 0 \end{bmatrix} \cdot \begin{bmatrix} \psi(A_{h1}) \cdots \psi(A_{hn}) \\ 0 \quad \cdots \quad 0 \\ \vdots \qquad\qquad \vdots \\ 0 \quad \cdots \quad 0 \end{bmatrix} \tilde{x}', \tilde{x}')$$

$$= \sum_{h=1}^{m} \| \begin{bmatrix} \psi(A_{h1}) \cdots \psi(A_{hn}) \\ 0 \quad \cdots \quad 0 \\ \vdots \qquad\qquad \vdots \\ 0 \quad \cdots \quad 0 \end{bmatrix} \tilde{x}' \|^2.$$

Thus $\| \sum_{j=1}^{n} \psi(A_{hj}) x_j' \| < \delta$; $h = 1, \cdots, m$. Let $\mathcal{V}'$ be the subspace of $\mathcal{H}'$ generated by $x_1', \cdots, x_n'$, and let $y_1', \cdots, y_n'$ be an orthonormal basis for $\mathcal{V}'$. Denote by S the linear transformation determined by $Sy_i' = x_i'$. We assert that $SS^* = \langle \tilde{x}', \tilde{x}' \rangle = \sum_{i=1}^{n} (\ , x_i') x_i'$. Indeed, $(S^* x_i', y_j') = (x_i', Sy_j') = (x_i', x_j')$, so that $S^* x_i' = \sum_{j=1}^{n} (x_i', x_j') y_j'$ and

$$SS^* x_i' = \sum_{j=1}^{n} (x_i', x_j') Sy_j' = \sum_{j=1}^{n} (x_i', x_j') x_j' = \langle \tilde{x}', \tilde{x}' \rangle (x_i').$$

Thus, since $x_1', \cdots, x_n'$ span $\mathcal{V}'$, $SS^* = \langle \tilde{x}', \tilde{x}' \rangle$ as asserted. It follows that SS^* has its spectrum in $[k^2, K^2]$ from which, $S^{-1*} S^{-1}$ has its spectrum in $[1/K^2, 1/k^2]$, so that $\| S \| \leqq K$ and $\| S^{-1} \| \leqq 1/k$. Let $S^* y_j' = \sum_{i=1}^{n} \alpha_{ij} y_i'$, so that the matrix of the transformation S^* relative to orthonormal basis $\{y_i'\}$ is (α_{ij}). In $\mathcal{V}$ define $y_j'' = \sum_{i=1}^{n} \alpha_{ij} x_i$. We set up a unitary transformation between $\mathcal{V}$ and $\mathcal{V}'$ by means of the map $x_i \rightarrow y_i'$. Under this map, we see that $y_j'' \rightarrow S^* y_j'$, so that

$$\begin{aligned}
&| (\sum_{j=1}^{n} y_j'' \otimes y_j'', \phi(U_i) y_i \otimes \phi(U_i) y_i - y_i \otimes y_i) | \\
&\quad = | (\sum_{j=1}^{n} S^* y_j' \otimes S^* y_j', [\sum_{h=1}^{n} \beta_{ih}' y_h'] \otimes [\sum_{h=1}^{n} \beta_{ih}' y_h'] - [\sum_{h=1}^{n} \beta_{ih} y_h'] \otimes [\sum_{h=1}^{n} \beta_{ih} y_h']) | \\
&\quad = | \sum_{j=1}^{n} (S^* y_j', \sum_{h=1}^{n} \beta_{ih}' y_h') (\sum_{h=1}^{n} \beta_{ih}' y_h', S^* y_j') \\
&\qquad - \sum_{j=1}^{n} (S^* y_j', \sum_{h=1}^{n} \beta_{ih} y_h') (\sum_{h=1}^{n} \beta_{ih} y_h', S^* y_j') | \\
&\quad = | \sum_{j=1}^{n} \{ | (y_j', \sum_{h=1}^{n} \beta_{ih}' x_h') |^2 - | (y_j', \sum_{h=1}^{n} \beta_{ih} x_h') |^2 \} | \\
&\quad = | \, \| \sum_{h=1}^{n} \beta_{ih}' x_h' \|^2 - \| \sum_{h=1}^{n} \beta_{ih} x_h' \|^2 | = | \, \| \sum_{h=1}^{n} \beta_{ih}' x_h' \|^2 - \| \psi(U_i) (\sum_{h=1}^{n} \beta_{ih} x_h') \|^2 | \\
&\quad \leqq \| \sum_{h=1}^{n} (\psi(U_i) \beta_{ih} x_h' - \beta_{ih}' x_h') \| \cdot [\| \sum_{h=1}^{n} \beta_{ih}' x_h' \| + \| \sum_{h=1}^{n} \psi(U_i) \beta_{ih} x_h' \|] \\
&\qquad \leqq \| \sum_{h=1}^{n} \psi(A_{ih}) x_h' \| (2n\beta K) \leqq 2n\beta K\delta,
\end{aligned}$$

where $\beta = \max \{ | \beta_{ih} |, | \beta_{ih}' |;\ i = 1, \cdots, m;\ h = 1, \cdots, n \}$.

In connection with the above inequality, note that we have proved that $\| x_i' \| \leqq K$, and that $\psi(U_i)$ is a unitary operator on $\mathcal{H}'$ since ψ is a * representation (with $\psi(I)=I$). We now specify the choice of δ as $1/2n\beta Kr$ (all the constants that appear in this choice were determined before the introduction of δ). Our inequality becomes then:

$$|(\sum_{j=1}^{n} y_j''\otimes y_j'', \phi(U_i)y_i\otimes\phi(U_i)y_i - y_i\otimes y_i)| \leqq 1/r; \qquad i=1,\cdots,m.$$

We write $y_j(r)$ for y_j'' to indicate the dependence of the y_j'' upon r. Observe that $y_j(r)$ lies in the sphere of radius K and outside the sphere of radius k, center at 0, in $\mathcal{V}$ since $y_j(r)$ is the image of S^*y_j' under our unitary map between $\mathcal{V}$ and $\mathcal{V}'$ (and $\|y_j'\|=1$, $\|S^*\|\leqq K$, $\|S^{*-1}\|\leqq 1/k$). By compactness, one can choose a subsequence $\{r_h\}$ of r's such that $\lim_h y_j(r_h)=z_j$; $j=1,\cdots,n$. Clearly

$$(\sum_{j=1}^{n} z_j\otimes z_j, \phi(U_i)y_i\otimes\phi(U_i)y_i - y_i\otimes y_i)=0; \qquad i=1,\cdots,m,$$

and the z_j lie between the spheres of radii k and K with center at 0 in $\mathcal{V}$. We consider the norm $\|\ \|'$ induced on $\mathcal{V}$ by means of the definition:

$$(\|x\|')^2=(\sum_{j=1}^{n} z_j\otimes z_j, x\otimes x)=\sum_{j=1}^{n}|(z_j,x)|^2.$$

We have just proved that $\sum_{j=1}^{n} z_j\otimes z_j$ is orthogonal to $\mathcal{E}\otimes$ so that $\|\phi(U)y\|'=\|y\|'$ if both $\phi(U)y$ and y are in $\mathcal{V}$. Thus, if $\phi(U)y$ and $\phi(V)y=\phi(VU^*)\phi(U)y$ are in $\mathcal{V}$ then $\|\phi(U)y\|'=\|\phi(V)y\|'$. For x an arbitrary vector in $\mathcal{V}$, we have $(\|x\|')^2=$

$$\lim_h(\sum_{j=1}^{n} y_j(r_h)\otimes y_j(r_h), x\otimes x)=\lim_h(\sum_{j=1}^{n} S_h^*y_j'(h)\otimes S_h^*y_j'(h), x'(h)\otimes x'(h))$$

$$=\lim_h\sum_{j=1}^{n}|(y_j'(h), S_hx'(h))|^2=\lim_h\|S_hx'(h)\|^2,$$

where $y_j'(h)$ are the vectors corresponding to y_j' in the foregoing discussion (with r_h now replacing r) S_h is the S of that discussion and $x'(h)$ is the image of x under the unitary map between $\mathcal{V}$ and $\mathcal{V}_h'$ of the present discussion. Now

$$\lim_h\|S_hx'(h)\|^2\leqq\lim_h K^2\|x'(h)\|^2=\lim_h K^2\|x\|^2=K^2\|x\|^2,$$

$$\lim_h\|S_hx'(h)\|^2\geqq\lim_h k^2\|x'(h)\|^2=\lim_h k^2\|x\|^2=k^2\|x\|^2,$$

so that $k\|x\|\leqq\|x\|'\leqq K\|x\|$.

If we take as partition classes in $\mathcal{H}$ the sets $\{\phi(U)x\colon U$ a unitary operator in $\mathfrak{A}\}$, we arive at a situation satisfying the hypotheses of Lemma 3, so that $\mathcal{H}$ admits a Hilbert space norm in which $\phi(U)$ is unitary for U a unitary operator in $\mathfrak{A}$ (this norm equivalent to the original norm with constants k, K). Thus, as at the end of the proof of Theorem 1, we can find an operator P with $P^{-1}\phi(U)P$ unitary for each unitary operator U in $\mathfrak{A}$ and $\| P \|$, $\| P^{-1} \|$ do not exceed $\max(K, 1/k)$. It now follows, as in the proof of Theorem 7, that the representation $A \rightarrow P^{-1}\phi(A)P$ is a * representation of $\mathfrak{A}$. Writing the polar decomposition HU, U unitary, $H = (PP^*)^{\frac{1}{2}}$ for P, we have $A \rightarrow H^{-1}\phi(A)H$ is a * representation of $\mathfrak{A}$ with H positive and having spectrum in $[k, K]$.

3. Concluding remarks. The discussion preceding Theorem 8 and Definition 3, concerning the approximate duplication of relations draws very heavily upon the fact that the initial algebra is a C^*-algebra (in particular, is uniformly closed) for the fact that an invertible operator in the algebra has its inverse in the algebra. On the other hand, Definition 3 and Theorem 8 apply as they stand to self-adjoint (not necessarily closed) algebras (although they are not stated this way). It follows immediately from this that:

COROLLARY 9. *A representation of a group by bounded operators on a Hilbert space is similar to a unitary representation if and only if the extension of this representation to the (finite, translation) group algebra (acting on* L_2 *of the group) has a self-adjoint cover.*

Despite the applicability of Definition 3 and Theorem 8 to self-adjoint algebras which are not uniformly closed, it should not be felt that the general conjecture about operator algebras has application to the non-closed, self-adjoint algebras. That is, examples are easily constructed of algebras which are not similar to self-adjoint algebras but are algebraically isomorphic to non-closed, self-adjoint algebras (not the continuous image, of course). In fact, let $x_1, x_2, \cdots$ be a sequence of linearly independent unit vectors which tend (strongly) to x and which span the Hilbert space $\mathcal{H}$. Let $\mathfrak{A}$ be the algebra of bounded operators A on $\mathcal{H}$ which have the form $Ax_i = a_i x_i$ for some sequence $\{a_i\}$ of complex numbers, and let $\mathcal{S}$ be the set of sequences which arise in this manner ($\mathcal{S}$ contains all sequences which have only a finite number of non-zero terms). Let $y_1, y_2, \cdots$ be an orthonormal basis for $\mathcal{H}$ and let $\mathfrak{A}'$ be the algebra of operators B of the form $By_i = a_i y_i$ where $\{a_i\}$ is in $\mathcal{S}$. Then $\mathfrak{A}'$ is a self-adjoint algebra containing I, for $B^* y_i = \bar{a}_i y_i$ and $\{\bar{a}_i\}$ is in $\mathcal{S}$ if and only if $\{a_i\}$ is in $\mathcal{S}$ (the x_i's being so chosen that the transformation $\alpha_1x_1 + \cdots + \alpha_nx_n \to \bar{\alpha}_1x_1 + \cdots + \bar{\alpha}_nx_n$ is bounded). Moreover the map $A \to B$ of $\mathfrak{A}$ onto $\mathfrak{A}'$ where $Ax_i = \alpha_ix_i$ and $By_i = \alpha_iy_i$ is an algebraic isomorphism (which is continuous, since $\| A \| \geqq \sup_i | \alpha_i | = \| B \|$).

For each invertible operator P and each operator A in $\mathfrak{A}$, the operator $P^{-1}AP$ has $P^{-1}x_i$ as eigenvectors and these converge to $P^{-1}x$. Now the algebra $\mathfrak{A}$ is commutative (as is $P^{-1}\mathfrak{A}P$) so that, if $P^{-1}\mathfrak{A}P$ is self-adjoint then $P^{-1}AP$ is normal for each A in $\mathfrak{A}$. Given $i \neq j$ we can easily find a sequence $\{\alpha_p\}$ in $\mathcal{S}$ with $\alpha_i \neq \alpha_j$ (let $\alpha_i = 1$, $\alpha_p = 0$ for $p \neq i$). Let A be the operator in $\mathfrak{A}$ with sequence $\{\alpha_p\}$. If $P^{-1}AP$ is normal then $P^{-1}x_i$ and $P^{-1}x_j$ are orthogonal. Thus if $P^{-1}\mathfrak{A}P$ is self-adjoint it follows that $P^{-1}x_i$, $i = 1, 2, \cdots$ is a set of mutually orthogonal vectors, which we have just seen cannot be the case.

We commented briefly, in the introduction, on the topological difficulty present in the infinite-dimensional case concerning the geometrical interpretation of semi-simplicity. By making suitable corrections for this difficulty, one arrives at a geometrical condition which might suffice for an algebra of operators to be similar to a self-adjoint algebra of operators. The conjecture obtained is quite natural in that it corrects for all the immediately visible difficulties which occur in passing from the finite to the infinite-dimensional case. For the moment, we specifically avoid describing the topology in which the operator algebra in question is closed.

Let $\mathfrak{A}$ be an algebra of operators on a Hilbert space with the property that there exists a positive δ such that if $\mathcal{V}$ is a closed subspace (*setwise*) *invariant under the operators of $\mathfrak{A}$ then there exists a complementary closed invariant subspace $\mathcal{W}$* (i.e., *$\mathcal{V} + \mathcal{W}$ is the whole space and $\mathcal{V} \cap \mathcal{W} = (0)$*) *which makes an angle greater than δ with $\mathcal{V}$. Is $\mathfrak{A}$ similar to a self-adjoint algebra?*

Note that since the angle between $\mathcal{W}$ and $\mathcal{V}$ is assumed to be positive their linear sum is closed. Let us assume that the answer to this question is yes (with any closure assumption on $\mathfrak{A}$) and that A is a bounded operator on the Hilbert space $\mathcal{H}$ with no closed invariant subspaces other than (0) and $\mathcal{H}$. Let $\mathfrak{A}$ be the (commutative) algebra generated by A and the identity operator (the closure taken in the appropriate topology). Since $\mathfrak{A}$ has no closed non-trivial invariant subspaces, the hypothesis is vacuously satisfied and there exists an operator P such that $P^{-1}\mathfrak{A}P$ is self-adjoint. Since $P^{-1}\mathfrak{A}P$ is commutative, it consists of normal operators. In particular $P^{-1}AP$ is normal and has an abundance of non-trivial, closed, invariant subspaces. If $\mathcal{V}$ is such a subspace then $P\mathcal{V}$ is non-trivial, closed and invariant under

A—a contradiction. Again, if $\mathfrak{A}$ is an irreducible algebra of operators then the hypothesis are trivially satisfied, and an affirmative answer to the question would imply that $\mathfrak{A}$ is similar to a self-adjoint algebra. Making use of this remark, we can answer the question in the uniformly closed case negatively. Our own approach to this counter-example rested upon producing a uniformly closed irreducible operator algebra containing an invertible operator whose inverse didn't lie in the operator algebra (note that this can't occur in an algebra which is similar to a C^*-algebra). A much more cogent device was suggested to us by I. Kaplansky. Using the completely continuous operators as a basic irreducible set of operators, build a closed operator algebra over it whose quotient by the completely continuous operators is a (finite-dimensional), non-semi-simple algebra. The larger algebra is not even the isomorphic image of a C^*-algebra, for a quotient algebra of a C^*-algebra is again a C^*-algebra [8] and therefore semi-simple. A concrete example is obtained by taking as our algebra the algebra generated by the completely continuous operators, the identity operator, and a nilpotent operator of index two (say a partial isometry between an infinite-dimensional subspace and its orthogonal complement).

In conclusion, we note the simple fact that a representation of a group by uniformly bounded operators each of which is normal is itself a unitary representation. In fact, an invertible normal operator all of whose powers form a set which is uniformly bounded in norm must have its spectrum on the unit circle and is therefore unitary.

Added in proof (June 1, 1955): In a recent note, (*Proceedings of the National Academy of Sciences*, vol. 41 (1955)) F. Mautner and L. Ehrenpreis announce that the group question has a negative answer, i. e., they produce a group and a bounded representation of it which is not similar to a unitary representation. Presumably, then, the "distortion continuity" condition of Theorem 8 cannot be removed. Restricting attention to relations involving n or fewer vectors, we can discuss representations satisfying an "n-distortion continuity" condition—the boundedness of a group representation (or continuity of a C^*-algebra representation) amounts to "1-distortion continuity." We feel that there are groups and representations of them which have n but not $n+1$ distortion continuity.

COLUMBIA UNIVERSITY.

BIBLIOGRAPHY.

[1] M. M. Day, "Means for the bounded functions and ergodicity of the bounded representations of semi-groups," *Transactions of the American Mathematical Society*, vol. 69 (1950), pp. 276-291.

[2] J. Dixmier, "Les moyennes invariantes dans les semi-groupes et leurs applications," *Acta Szeged*, vol. 12 (1950), pp. 213-227.

[3] F. J. Murray and J. von Neumann, "On rings of operators," *Annals of Mathematics*, vol. 37 (1936), pp. 116-229.

[4] B. Sz. Nagy, "On uniformly bounded linear transformations in Hilbert space," *Acta Szeged*, vol. 11 (1947), pp. 152-157.

[5] M. Nakamura and z. Takeda, "Group representations and Banach limits," *Tôhoku Mathematical Journal*, vol. 3 (1951), pp. 132-135.

[6] G. Mackey, *Commutative Banach algebras*, Mimeographed lecture notes, Harvard, 1952.

[7] F. Peter and H. Weyl, "Die Vollstandigkeit der primitiven Darstellungen einer geschlossenen kontinuierlichen Gruppe," *Mathematische Annalen*, vol. 97 (1927), pp. 737-755.

[8] I. Segal, "Two-sided ideals in operator algebras," *Annals of Mathematics*, vol. 50 (1949), pp. 856-865.

CENTRAL LIBRARY UNIVERSITY

ON THE GENERATORS OF AN IDEAL, WITH AN APPLICATION TO THE GEOMETRY OF NUMBERS IN UNITARY SPACE U_2.*

By K. Rogers.

---

1. In the course of joint work with H. P. F. Swinnerton-Dyer on the geometry of numbers in unitary space, U_n denoting n-dimensional vector-space over the field of complex numbers, the author proved some results on algebraic numbers. These are given below and applied to the general theory of lattices in U_2, with special reference to the problem of two linear forms with complex coefficients and variables integers of an algebraic number field.

Let k be the field of rationals, θ a number algebraic over k, and $k(\theta)$ the field obtained by adjunction of θ. 'Integer' will mean 'integer of $k(\theta)$' unless otherwise specified. Every ideal $\mathfrak{a}$ can be expressed as (p, q), where p, q are integers which we shall call a *generating pair.* Also, every ideal has a finite *basis* as module over the rational integers, and in particular there is a basis of two elements when θ is quadratic over k. We prove the following results:

Theorem 1. *If* $(p, q) = (r, s) = \mathfrak{a}$, *then for any integer* u *of* $k(\theta)$ *there exists an integral* 2×2 *matrix* A *of determinant* u, *such that* $(p, q)A = (r, s)$.

When $k(\theta)$ is quadratic, we have in particular:

Corollary. *If* (a, b) *is a basis of* $\mathfrak{a}$, *then all generating pairs of* $\mathfrak{a}$ *are obtained from* (a, b) *by transformations with coefficients integral and determinant a unit.*

Calling this type of transformation 'integral unimodular,' we also prove:

Theorem 2. *If* $\mathfrak{M}(e)$ *is the set of all* 2×2 *matrices* A *with integral elements and* $\text{Norm}(\det A) \leqq e$, *then there exists a finite subset* $\{U_1, \cdots, U_h\}$ *of* $\mathfrak{M}(e)$ *such that to any* $A \varepsilon \mathfrak{M}(e)$ *there corresponds a* U_m *and a unimodular integral matrix* T *such that* $A = U_m T$.

The referee pointed out that the case $u = 1$ of Theorem 1 is *Satz* 53 in Hilbert's *Zahlbericht* and is due to Hurwitz [1]. Also, Swinnerton-Dyer

---

* Received October 8, 1954; revised May 27, 1955.

and I have proved Theorem 2 for $n \times n$ matrices by considering lattices in U_n, but the proofs of both theorems given here are needed for the application to linear forms.

2. Proof of Theorem 1. Since $\mathfrak{a} = (p, q) = (r, s)$, there exist integers a, b, c, d such that $r = ap + bq$, $s = cp + dq$. Hence,

$$r = (a + f)p + (b - g)q, \qquad s = (c + f')p + (d - g')q,$$

where f, g, f', g' are integers such that

$$fp = gq, \qquad f'p = g'q. \tag{2.1}$$

Clearly, g and g' are any integers such that qg/p and qg'/p are also integers, that is, are any elements of the ideal $(p)/(q, p)$. It remains to choose solutions of (2.1) such that $(a + f)(d - g') - (b - g)(c + f') = u$. Denoting this determinant by Δ, setting $ad - bc = \Delta_0$, and noting that $fg' = f'g$, we have

$$\Delta = \Delta_0 - ag' + df + cg - bf',$$

and hence

$$\begin{aligned} p\Delta &= p\Delta_0 - apg' + dpf + cpg - bpf' \\ &= p\Delta_0 - apg' + dqg + cpg - bqg' \\ &= p\Delta_0 + gs - g'r. \end{aligned} \tag{2.2}$$

Now $gs - g'r$ takes any values in the ideal $(p)(r, s)/(q, p) = (p)$, and in particular the value $p(u - \Delta_0)$. With this value in (2.2), we have $\Delta = u$, as required.

3. Proof of Theorem 2. By the bound on Norm$(\det A)$, the number of different ideals $(\det A)$ for $A \varepsilon \mathfrak{M}(e)$ is finite. We need therefore only prove the theorem for matrices of determinant E, say, where Norm$(E) \leqq e$, since unit factors can be absorbed into the unimodular matrices. Let A denote such a matrix, and write

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Since $ad - bc = E$, the ideal $\mathfrak{a} = (a, b)$ must divide E and is therefore one of a finite number of ideals, finiteness depending on E and hence on e. For each of these ideals we can choose generating pairs, and then, by Theorem 1, we can multiply A on the right by a unimodular integral matrix such that the new first row is one of a finite number. If it happens that the row-ideal is principal, then we can find a unimodular T such that

$$AT=\begin{pmatrix} a' & 0 \\ b' & c' \end{pmatrix},$$

in which case $a'c'=E$, and so apart from units a' and c' belong to a finite set of integers depending on e. Then we can reduce b' modulo c', this corresponding to multiplication on the right by a unimodular matrix, by substracting multiples of the second column from the first. In this way we should have found a unimodular integral matrix T such that AT is a matrix whose elements are integers out of a finite set depending only on e. Since T has an integral unimodular inverse, the theorem is proved in the particular case just considered. We now suppose that the ideal $\mathfrak{a}$ is non-principal, so that if

$$AT=\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix},$$

where a' and b' belong to a finite set of integers, then $a'b'$ is not zero. We now consider what unimodular transformations can be applied so as to leave the first row invariant and to vary the element c' modulo some integer which is bounded in terms of e. We have

$$\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}\begin{pmatrix} 1+f_1 & f_2 \\ -g_1 & 1-g_2 \end{pmatrix}=\begin{pmatrix} a' & b' \\ c'+c'f_1-d'g_1 & \cdots \end{pmatrix}$$

where we require that

$$a'f_i=b'g_i \qquad (i=1,2). \tag{3.1}$$

We wish also to satisfy $\begin{vmatrix} 1+f_1 & f_2 \\ -g_1 & 1-g_2 \end{vmatrix}=1$, or

$$f_1=g_2. \tag{3.2}$$

We can certainly take $f_1=g_2=ua'b'$, where u is any integer, and then $f_2=ub'^2$, $g_1=ua'^2$. In this case

$$c'f_1-d'g_1=ua'(c'b'-a'd')=-a'uE,$$

so that the element c' can be varied modulo $-a'E$. In certain special cases, which are important in applications, we can vary c' modulo E. These are the cases when a' divides b'^2, when we can take the set of values

$$f_1=g_2=ub', \qquad g_1=ua', \qquad f_2=ub'^2/a',$$

which then give $c'f_1-d'g_1=u(c'b'-d'a')=-uE$. An example of the second possibility is the ideal $(2, 1 + i\sqrt{5})$ in the ring of integers of $k(i\sqrt{5})$. In any case, we have now found a unimodular integral matrix T_1 such that

$$AT_1 = \begin{pmatrix} a' & b' \\ c'' & d'' \end{pmatrix},$$

where now a', b' and c'' are bounded in terms of e. But since a' is not zero we can write $d'' = (E + b'c'')/a'$, and so d'' also has only a finite number of possible values. This completes the proof of Theorem 2.

4. Applications of these results. We require some definitions analogous to those of Mahler [3]. A function $f(\boldsymbol{x}) = f(x_1, \cdots, x_n)$ is called a *distance function* in U_n if

1) $f(\boldsymbol{x})$ is continuous and real-valued for complex x.

2) $f(\boldsymbol{x}) \geqq 0$ for all $\boldsymbol{x} \in U_n$ and $\neq 0$ at at least one point.

3) $f(t\boldsymbol{x}) = |t| f(\boldsymbol{x})$, for all $\boldsymbol{x} \in U_n$ and all complex t.

The set K of all points satisfying $f(\boldsymbol{x}) \leqq 1$, where $f(\boldsymbol{x})$ is a distance function, is called a *star-body*. A *lattice in* $k(\theta)$ is defined as the set Λ of all points $\boldsymbol{x} = A\boldsymbol{u}$, $(\det A \neq 0)$ where $\boldsymbol{u}$ runs through all $n \times 1$ column-vectors with elements integers of $k(\theta)$. Since $T\boldsymbol{u}$, for unimodular integral T, runs through all such column-vectors, the matrix AT also defines the same lattice. Since $|\det(AT)| = |\det A|$, this is a well-defined positive number which is called the determinant of the lattice, $d(\Lambda)$. A lattice Λ is said to be *admissible* for a star-body K if K contains no point of Λ in its interior other than the origin. Since the origin is an inner point of K, it is not difficult to show that there exists a positive number c, depending only on K, such that all lattices with $d(\Lambda) < c$ are inadmissible for K. We define the *critical determinant* $\Delta(K)$ as the lower bound of the determinants of admissible lattices, with the convention that $\Delta(K) = \infty$ if there is no admissible lattice. The preceding remark implies that for star-bodies $\Delta(K) > 0$. A lattice of determinant $\Delta(K)$ which is admissible for K is called a *critical* lattice of K. The central problem of the geometry of numbers is to find $\Delta(K)$ for various types of star-bodies and to find all possible critical lattices. Our object will be to prove that for star-bodies in U_2 with finite critical determinant, and for $k(\theta)$ an imaginary quadratic field over k, critical lattices always exist. For this we define a *bounded sequence of lattices* $\{\Lambda_r\}$ as a sequence for which there exist constants c_1, c_2 such that $d(\Lambda_r) \leqq c_1$ and $|\boldsymbol{x}| = \sqrt{\boldsymbol{x}^*\boldsymbol{x}} \geqq c_2$ for all $\boldsymbol{x} \in \Lambda_r$, $\boldsymbol{x} \neq 0$, $r = 1, 2, \cdots$.

LEMMA 1. *If $k(\theta)$ is complex quadratic, then every bounded infinite sequence of lattices in U_2 contains a convergent infinite subsequence.*

Proof. We begin by showing that Λ_r has independent points P_r, Q_r bounded in terms of c_1 and c_2, and for this we use a theorem of H. Weyl [6]. If $f(\boldsymbol{x})$ is a distance-function in U_n, let V denote the content of the star-body corresponding to that given by $f(\boldsymbol{x}) \leqq 1$ obtained by taking the real and imaginary parts of each co-ordinate as a co-ordinate in Euclidean $2n$-space. Let M_1 be the smallest positive number such that the region $f(\boldsymbol{x}) \leqq M_1$ has on its boundary a point of a given lattice Λ, say $\boldsymbol{x}^1$, and define M_k inductively as the least M such that $f(\boldsymbol{x}) \leqq M$ has on its boundary a point $\boldsymbol{x}^k$ of Λ which is independent of $\boldsymbol{x}^1, \cdots, \boldsymbol{x}^{k-1}$. By applying a result of Minkowski, Weyl showed that

$$M_1{}^2 \cdots M_n{}^2 V \leqq (2d^{\frac{1}{2}})^n \{d(\Lambda)\}^2,$$

where $-d$ is the discriminant of the quadratic field $k(\theta)$. In the present case we take

$$f(x) = \sqrt{\boldsymbol{x}^*\boldsymbol{x}} = \sqrt{(|x_1|^2 + |x_2|^2)},$$

so that V is the content of the 4-dimensional sphere of unit radius. Since $M_1 \geqq c_2$, we have

$$M_1{}^2 \leqq M_2{}^2 \leqq 4d\{d(\Lambda)\}^2 V^{-1} c_2{}^{-2},$$

and so that points $P_r = \boldsymbol{x}_r{}^1$, $Q_r = \boldsymbol{x}_r{}^2$ are of non-zero determinant and are bounded independently of r. Taking a matrix A_r which represents the lattice, we suppose that $\boldsymbol{x}_r{}^1 = A_r \boldsymbol{u}_r{}^1$, $\boldsymbol{x}_r{}^2 = A_r \boldsymbol{u}_r{}^2$, noting that if we take the matrix $A_r T^{-1}$ to represent the lattice, then the corresponding integer vectors are $T\boldsymbol{u}_r{}^1$ and $T\boldsymbol{u}_r{}^2$. Since

$$\det(\boldsymbol{x}_r{}^1, \boldsymbol{x}_r{}^2) = \det A_r \cdot \det(\boldsymbol{u}_r{}^1, \boldsymbol{u}_r{}^2),$$

and since $|\det A_r|$ has a lower bound independent of r, owing to the fact that Λ_r is admissible for the sphere of radius c_2, we have an upper bound for $|\det(\boldsymbol{u}_r{}^1, \boldsymbol{u}_r{}^2)|$ which is independent of r. By Theorem 2 applied to transposed matrices, we conclude that there exists a finite set of integral matrices $\{U_1, \cdots, U_m\}$ such that for each r there exists a unimodular integral matrix T_r with $T_r(\boldsymbol{u}_r{}^1, \boldsymbol{u}_r{}^2) = U_i$ for some i, $1 \leqq i \leqq m$. Thus we can choose a subsequence of $\{\Lambda_r\}$ such that for each lattice there is a basis for which the points P, Q have co-ordinates independent of r. We take this subsequence and rename it $\{\Lambda_r\}$ for convenience. Since the values of $d(\Lambda_r)$ are bounded, we can choose an infinite subsequence of $\{\Lambda_r\}$ such that $d(\Lambda_r)$ tends to some positive number, say Δ. Renaming this subsequence $\{\Lambda_r\}$, for convenience of notation, we now have a sequence of lattices $\{\Lambda_r\}$ and point-pairs P_r, Q_r, such that $d(\Lambda_r) \rightarrow \Delta$, P_r and Q_r are bounded independently of r, and there is a matrix A_r representing Λ_r such that $(P_r, Q_r) = A_r U$, where U is an integral matrix independent of r. Since P_r and Q_r are bounded independently of r, we can choose a subsequence of $\{\Lambda_r\}$, which we rename once more, such that all the above conditions are satisfied and also $P_r \rightarrow P$, $Q_r \rightarrow Q$, as $r \rightarrow \infty$. Then, since U is fixed and non-singular, the fact that $A_r U$ tends to a limit as $r \rightarrow \infty$ implies that there exists a matrix A of determinant of modulus Δ, such that $A_r \rightarrow A$ as $r \rightarrow \infty$. Hence this subsequence of the original $\{\Lambda_r\}$ converges to the lattice Λ corresponding to A in the sense that there exists a basis of Λ (represented by the columns of A) to which a basis of a lattice of the sequence converges as $r \rightarrow \infty$. The lemma is proved.

THEOREM 3. *With respect to the complex quadratic field $k(\theta)$, every star-body in U_2 of "finite type" (that is, with finite $\Delta(K)$) possesses at least one critical lattice.*

Proof. By the definition of $\Delta(K)$, there exists a sequence $\{\Lambda_r\}$ of lattices admissible for K such that $d(\Lambda_r) \rightarrow \Delta(K)$ as $r \rightarrow \infty$. By discarding a finite number of these, we can suppose that $d(\Lambda_r) \leqq 2\Delta(K)$. As we remarked earlier, there exists a sphere round the origin which lies entirely in K; hence this and the previous inequality imply that the sequence $\{\Lambda_r\}$ is bounded. By the lemma there is a subsequence, which we rename $\{\Lambda_r\}$, which converges to a lattice Λ of determinant $\Delta(K)$. To prove that Λ is critical for K it remains only to prove that Λ is admissible for K. Taking bases (P_1, P_2) and (P_{r1}, P_{r2}) such that $(P_{r1}, P_{r2}) \rightarrow (P_1, P_2)$ as $r \rightarrow \infty$, suppose the point $Q = u_1 P_1 + u_2 P_2 \neq 0$ of Λ lies in the interior of K. Then the point $Q_r = u_1 P_{r1} + u_2 P_{r2} \neq 0$ of Λ_r tends to $u_1 P_1 + u_2 P_2$ as $r \rightarrow \infty$, and hence, since the distance function of K is continuous, the point Q_r lies in the interior of K for large enough r. This contradicts the fact that Λ_r is K-admissible, and hence Λ must be a critical lattice of K.

5. The above theorem has practical application, quite apart from its interest as a pure existence theorem. A lattice Λ is said to be *extremal* for K if in a sufficiently small 'neighbourhood of Λ' (defined in an obvious way) there are no K-admissible lattices Λ' such that $d(\Lambda') < d(\Lambda)$; clearly a critical lattice is extremal. Thus, if we have *necessary* conditions for an extremal lattice, and if these are satisfied by only a finite number of lattices, then since critical lattices exist we know that a lattice of smallest determinant is a critical lattice. This has been applied by the author to the problem of maximising the function

$$\min_{u,v} \{\max(|\alpha u + \beta v|, |\gamma u + \delta v|)\}$$

of complex numbers α, β, γ, δ such that $\alpha\delta-\beta\gamma=\Delta\neq 0$, the above minimum being taken over all integers u, v of $k(\theta)$ which are not both zero. With the help of an important lemma of Swinnerton-Dyer, the author found simple proofs for the cases $k(i)$ and $k(\rho)$, first considered by Minkowski [4]. We have now found the best possible result for $k(i\sqrt{5})$, when new problems arise from the class-number being 2. One has two columns of integers, the rows representing lattice points and generating either (1) or $(2, 1+i\sqrt{5})$. The determinants are bounded, and we wish to reduce the array by unimodular right-multiplication to a finite number of possible cases. The proof of Theorem 2 shows this is trivial if a row-ideal is (1), and otherwise one row can be taken as $(2, 1+i\sqrt{5})$, while the first element of the next row can be reduced modulo the determinant E. The details of this result will be published in a joint paper with Swinnerton-Dyer. This will develop general theory in U_n and will include a complete analogue of results of Korkine and Zolotareff [2] and Swinnerton-Dyer [5].

In conclusion, I thank the referee for simplifying the proof of Theorem 1 and for the reference to Hurwitz' paper.

THE GRADUATE COLLEGE,
PRINCETON, N. J.

REFERENCES.

[1] A. Hurwitz, "Die unimodularen Substitutionen in einem algebraischen Zahlenkörper," *Mathematische Werke* (Verlag Birkhäuser, 1933), Bd. II, 244-268.

[2] A. Korkine and G. Zolotareff, "Sur les formes quadratiques positives," *Mathematische Annalen*, vol. 11 (1877), pp. 242-292.

[3] K. Mahler, "On lattice-points in n-dimensional star-bodies," *Proc. Royal Soc.* A, vol. 187 (1946), pp. 151-187.

[4] H. Minkowski, *Diophantische Approximationen* (Leipzig, B. G. Teubner, 1907), Chapter 6.

[5] H. P. F. Swinnerton-Dyer, "Extremal lattices of convex bodies," *Proc. Camb. Phil. Soc.*, vol. 49 (1953), p. 161.

[6] H. Weyl, "Theory of reduction for arithmetical equivalence," *Transactions of the American Mathematical Society*, vol. 48 (1940), pp. 126-164 (particularly Theorem 1*, p. 140).

AN OUTLINE OF THE THEORY OF F_2 SERIES.*

By S. VERBLUNSKY.

1. *Introduction.* Let $A(z)$, $B(z)$ be relatively prime polynomials in z, with real coefficients, such that the equation

$$(1) \qquad A \sin \pi z - B \cos \pi z = 0$$

has all its roots real and simple. Then any function $f(x)$ which is L-integrable in an interval of length 2π, say $(\alpha, \alpha + 2\pi)$, has an expansion of "Fourier" type, in terms of cosines and sines of $\mu_k x$. Here, $\mu_1 < \mu_2 < \cdots$ are the positive roots of (1), $\mu_{-1} > \mu_{-2} > \cdots$ are the negative roots, and, if 0 is a root, then $\mu_0 = 0$. ([3], Theorem 1.) To fix the ideas, we shall consider the case in which the roots are symmetrical with respect to the origin. It is known that this happens if and only if one of A, B is odd, and the other is even. ([3], Section 5.) The "Fourier" expansion can then be written

$$(2) \qquad \tfrac{1}{2}a_0 + \sum_{k=1}^{\infty} a_k \cos \mu_k x + b_k \sin \mu_k x,$$

where $a_k - ib_k = 2\lambda_k \int_\alpha^{\alpha+2\pi} f(x) e^{-i\mu_k x}\, dx$, $(k = 1, 2, \cdots)$, and, if 0 is a root of (1), $a_0 = 2\lambda_0 \int_\alpha^{\alpha+2\pi} f(x)\, dx$, while, if 0 is not a root of (1), then $a_0 = 0$. The numbers λ_k are the residues, at $z = \mu_k$, of

$$Q(z) = \tfrac{1}{2}(A \cos \pi z + B \sin \pi z)/(A \sin \pi z - B \cos \pi z).$$

The object of this paper is to investigate the series (2) when the even one of A, B is of degree 2, and the odd one is of degree 1. If A is the even one, we shall call (2) the F_2A series of $f(x)$ in $(\alpha, \alpha + 2\pi)$. If B is the even one, we shall call (2) the F_2B series of $f(x)$ in $(\alpha, \alpha + 2\pi)$. It is no loss of generality to suppose that the leading coefficient of the even one is unity. Thus, for F_2A series we can write $A = z^2 + a$, $B = bz$, and for F_2B series, $A = az$, $B = z^2 + b$, and in both cases, $ab \neq 0$, since A, B are relatively prime.

The properties of an F_2A series depend on whether $b^2 + 4a$ is positive, zero, or negative. In this paper, we consider only the first case. Similarly

---

* Received January 18, 1955.

for F_2B series, we consider only the case in which a^2+4b is positive. On examining Theorems 3 and 4 of [3], we can make the following deductions. The points (a,b) such that, (i) $b^2+4a>0$ and (ii) such that the equation

$$(z^2+a)\sin\pi z-bz\cos\pi z=0$$

has all its roots real and simple, are the points below the curve

$$y=-x,\ x\geqq 0;\qquad y=-2\,|\,x\,|^{\frac{1}{2}},\ x<0.$$

The points (a,b) such that, (i) $a^2+4b>0$ and (ii) such that the equation $az\sin\pi z-(z^2+b)\cos\pi z=0$ has all its roots real and simple, are not so easily described. However, they include all the points (a,b) with $a<0$, $b>0$, $b^2+4a>0$.

We now mention some of the results proved below, and, for brevity, we refer only to F_2A series.

The coefficients of an F_2A series in $(\alpha,\alpha+2\pi)$ are connected by two linear equations (Theorem 1). It is convenient to have a name for the series

$$\sum_{2}^{\infty} a_k\cos\mu_k x+b_k\sin\mu_k x, \tag{3}$$

derived from the F_2A series (2) by omitting the initial terms. We call (3) a "restricted F_2A series" (RF_2A series) in $(\alpha,\alpha+2\pi)$. Given an RF_2A series (3) in $(\alpha,\alpha+2\pi)$, the "missing" terms, viz. $\frac{1}{2}a_0+a_1\cos\mu_1x+b_1\sin\mu_1x$ are, by Theorem 5, determined to within an arbitrary additive multiple of

$$-(a/b)\cdot\sin\pi\mu_1/\mu_1+\cos\mu_1(x-\alpha-\pi).$$

An RF_2A series in $(\alpha,\alpha+2\pi)$ is an RF_2A series in $(\beta,\beta+2\pi)$ for all β (Theorem 4). An RF_2A series is summable (C, 1) almost everywhere, and is the Fourier series of its Cesarò sum (in the sense of the theory of almost periodic functions), by Theorem 7. Theorems of the Young-Hausdorff type hold for F_2A series (Theorem 8). In enunciating such theorems, we must take account of the fact that there are two linear equations connecting the coefficients of an F_2A series. As an example, we quote the following special case of Theorem 8(a): Let c_ν, $\nu=2,3,\cdots$, be a sequence of complex numbers such that $\sum|\,c_\nu\,|^p<\infty$, where $1<p\leqq 2$. Let $1/p+1/q=1$. Then there is precisely one real function $f\,\varepsilon\,L^q(\alpha,\alpha+2\pi)$ which satisfies the conditions

$$0=\lambda_0\int_\alpha^{\alpha+2\pi}f(t)\,dt,\qquad c_\nu=\lambda_\nu\int_\alpha^{\alpha+2\pi}f(t)\,e^{-i\mu_\nu t}dt.\qquad(\nu=2,3,\cdots)$$

For this function,

$$\Big(\int_\alpha^{\alpha+2\pi}|\,f\,|^q\Big)^{1/q}\leqq\Big(2\sum_{1}^{\infty}|\,c_\nu\,|^p/\lambda_\nu^{\,p-1}\Big)^{1/p},\ \text{where }c_1=\lambda_1\int_\alpha^{\alpha+2\pi}f(t)\,e^{-i\mu_1 t}dt.$$

Finally, as an example of a uniqueness theorem, we have: If

$$\sum_{1}^{\infty} (a_k \cos \mu_k x + b_k \sin \mu_k x) \tag{4}$$

converges to zero for all x in the open interval $(-\pi, \pi)$, then (4) is obtained by termwise differentiation of the F_2A series of a constant in $(-\pi, \pi)$ (Theorem 9).

Sections 2-11. F_2A Series.

2. As mentioned in Section 1, we may suppose that

$$A = z^2 + a, \qquad B = bz, \qquad ab \neq 0. \tag{5}$$

Writing $A_k = A(\mu_k)$, etc., we have

$$\lambda_k = \tfrac{1}{2} \frac{A_k \cos \pi\mu_k + B_k \sin \pi\mu_k}{\pi(A_k \cos \pi\mu_k + B_k \sin \pi\mu_k) + A_k' \sin \pi\mu_k - B_k' \cos \pi\mu_k}.$$

But

$$A_k \sin \pi\mu_k = B_k \cos \pi\mu_k. \tag{6}$$

Hence

$$1/(2\lambda_k) = \pi + b(\mu_k^2 - a)/[(\mu_k^2 + a)^2 + \mu_k^2 b^2], \tag{7}$$

and, in particular,

$$1/(2\lambda_0) = \pi - b/a. \tag{8}$$

Let f be a real function of class $L(\alpha, \alpha + 2\pi)$. Write

$$c_k = \lambda_k \int_{\alpha}^{\alpha+2\pi} f(t) e^{-i\mu_k t} dt, \qquad (k = 0, \pm 1, \cdots),$$

so that $c_{-k} = \bar{c}_k$. Then the F_2A series of $f(x)$ in $(\alpha, \alpha + 2\pi)$ can be written $\sum c_k e^{i\mu_k x}$. The relation between this series and the function f in the interval $(\alpha, \alpha + 2\pi)$, will be denoted by

$$f(x) \sim \sum c_k e^{i\mu_k x}, \qquad (\alpha, \alpha + 2\pi). \tag{9}$$

If δ is a real number, then

$$f(x - \delta) \sim \sum c_k e^{-i\mu_k \delta} e^{i\mu_k x}, \qquad (\alpha + \delta, \alpha + \delta + 2\pi), \tag{10}$$

and

$$f(-x) \sim \sum c_{-k} e^{i\mu_k x}, \qquad (-\alpha - 2\pi, -\alpha). \tag{11}$$

3. As regards convergence, F_2A series behave like Fourier series. This is a special case of a known theorem ([3], Theorem 1), and will be enunciated as a lemma.

LEMMA 1. *If* $s_n(x)$ *denotes the n-th partial sum of* (9), *viz.,*

$$s_n(x) = \sum_{-n}^{n} c_k e^{i\mu_k x},$$

and $\sigma_n(x)$ *denotes the n-th partial sum of the Fourier series of* $f(x)$, *viz.,*

$$\sigma_n(x) = \sum_{-n}^{n} \frac{1}{2\pi} \int_{\alpha}^{\alpha+2\pi} f(t) e^{-ik(t-x)} dt,$$

then $s_n(x) - \sigma_n(x) \to 0$ *uniformly in* $(\alpha, \alpha + 2\pi)$ *as* $n \to \infty$. *In particular,* $s_n(\alpha + 2\pi) - s_n(\alpha) \to 0$ *as* $n \to \infty$.

In virtue of the last clause of Lemma 1,

$$\sum{}' c_k e^{i\mu_k \alpha} (e^{2\pi i \mu_k} - 1) = 0, \tag{12}$$

where here, and in what follows, the accent denotes that the term with $k = 0$ is absent.

The equation (1) can be wirtten in the form $e^{2\pi i z} = (A + iB)/(A - iB)$. We have $A - iB = z^2 - ibz + a = (z - im_1)(z - im_2)$. To keep our arguments as simple as possible, we shall suppose that $b^2 + 4a > 0$. Thus, m_1, m_2 are real and different. Then

$$e^{2\pi i \mu_k} - 1 = 2iB_k/(A_k - iB_k) = 2ib\mu_k/[(\mu_k - im_1)(\mu_k - im_2)]. \tag{13}$$

We can now prove that the numbers c_k in (9), which we may call the F_2A constants of f in $(\alpha, \alpha + 2\pi)$, are connected by two linear relations, independent of f.

THEOREM 1. *For any* F_2A *series* (9), *we have*

$$\sum{}' c_k e^{i\mu_k \alpha} \mu_k/[(\mu_k - im_1)(\mu_k - im_2)] = 0, \tag{14}$$

and

$$c_0/a + \sum{}' c_k e^{i\mu_k \alpha}/[(\mu_k - im_1)(\mu_k - im_2)] = 0, \tag{15}$$

or, what is equivalent, the two equations

$$\sum c_k e^{i\mu_k \alpha}/(\mu_k - im_1) = 0, \qquad \sum c_k e^{i\mu_k \alpha}/(\mu_k - im_2) = 0. \tag{16}$$

Proof. Equation (14) follows from (12) on substituting the value of $e^{2\pi i \mu_k} - 1$ from (13). To obtain (15), consider the F_2A series of the function defined as equal to 1 in $(\alpha, \alpha + 2\pi)$. If

(17) $$1 \sim \sum c_k^* e^{i\mu_k x}, \qquad (\alpha, \alpha + 2\pi),$$

then

(18) $$c_0^* = 2\pi\lambda_0, \quad c_k^* = (\lambda_k/-i\mu_k)\, e^{-i\mu_k\alpha}(e^{-2\pi i\mu_k} - 1), \qquad (k \neq 0).$$

On account of the uniform equiconvergence of an F_2A series and a Fourier series, we can multiply (17) by $f(x)$ and integrate term by term. We thus obtain

$$c_0/\lambda_0 = 2\pi c_0 + \sum{}' \,(c_{-k}/-i\mu_k)\, e^{-i\mu_k\alpha}(e^{-2\pi i\mu_k} - 1).$$

By (8) and (13), this becomes

$$2b/a \cdot c_0 + \sum{}' \, c_{-k}\, e^{-i\mu_k\alpha}\, 2b/[(\mu_k + im_1)(\mu_k + im_2)] = 0.$$

Divide by $2b$, and take conjugates. We obtain (15).

If we multiply (15) by $-im_2$ and add to (14), we obtain

$$-c_0 i m_2/a + \sum{}' \, c_k\, e^{i\mu_k\alpha}/(\mu_k - im_1) = 0.$$

But $a = -m_1 m_2$. Hence the first term is $c_0\, e^{i\mu_0\alpha}/(\mu_0 - im_1)$. This proves the first equation (16). The second equation is proved similarly. By reversing the argument, we see that equations (16) imply (14) and (15).

4. In order to decide when the trigonometric polynomial

$$\sum_{|k| \leqq n} c_k e^{-i\mu_k x}, \qquad (c_{-k} = \bar{c}_k),$$

is an F_2A series in $(\alpha, \alpha + 2\pi)$, we need to examine two special cases. We first observe that if (n, ν) are positive integers, then by (6),

(19) $$\int_{-\pi}^{\pi} \cos \mu_\nu x \cos \mu_n x \, dx = 2a/b \cdot \sin \pi\mu_\nu/\mu_\nu \cdot \sin \pi\mu_n/\mu_n, \qquad (\nu \neq n),$$

and

(20) $$\int_{-\pi}^{\pi} \sin \mu_\nu x \sin \mu_n x \, dx = -2/b \cdot \sin \pi\mu_\nu \sin \pi\mu_n, \qquad (\nu \neq n).$$

LEMMA 2. (i) *If ν is a positive integer, then*

(21) $$-(a/b) \cdot \sin \pi\mu_\nu/\mu_\nu + \cos \mu_\nu x$$

is its own F_2A series in $(-\pi, \pi)$.

(ii) *If $\nu, > 1$, is a positive integer, then*

(22) $$-\sin \mu_1 x/\sin \mu_1 \pi + \sin \mu_\nu x/\sin \mu_\nu \pi$$

is it own F_2A series in $(-\pi, \pi)$.

Proof. (i) The F_2A series of the even function (21) can be written $\frac{1}{2}a_0 + \sum_{1}^{\infty} a_n \cos \mu_n x$. If $n \neq 1, \nu$, then

$$a_n = 2\lambda_n \int_{-\pi}^{\pi} (-(a/b)\cdot \sin\, \pi\mu_\nu/\mu_\nu + \cos \mu_\nu x) \cos \mu_n x \, dx = 0,$$

by (19). Thus, the F_2A series of (21) is $\frac{1}{2}a_0 + a_\nu \cos \mu_\nu x$, and this converges to (21) in $(-\pi, \pi)$. Put $x = \pi/2\mu_\nu$. Then $\frac{1}{2}a_0 = -(a/b)\cdot \sin \pi\mu_\nu/\mu_\nu$. Hence also $a_\nu = 1$.

(ii) The proof is similar.

We can now solve the problem mentioned at the beginning of this section.

THEOREM 2. *In order that*

$$\sum_{|k| \leqq n} c_k e^{i\mu_k x}, \qquad (c_{-k} = \bar{c}_k), \tag{23}$$

shall be an F_2A series in $(\alpha, \alpha + 2\pi)$, it is necessary and sufficient that

$$\sum_{|k| \leqq n}{}' c_k e^{i\mu_k \alpha}(e^{2\pi i \mu_k} - 1) = 0, \tag{24}$$

and that

$$c_0/a + \sum_{|k| \leqq n}{}' c_k e^{i\mu_k \alpha}(e^{2\pi i \mu_k} - 1)/(2ib\mu_k) = 0; \tag{25}$$

or, what is equivalent, that

$$\sum_{|k| \leqq n} c_k e^{i\mu_k \alpha}/(\mu_k - im_1) = 0, \qquad \sum_{|k| \leqq n} c_k e^{i\mu_k \alpha}/(\mu_k - im_2) = 0. \tag{26}$$

Proof. That the conditions are necessary, follows from Theorem 1 on using (13). We further observe, on recalling the argument at the end of Section 3, that (24) and (25) are equivalent to (26).

To prove that the conditions (24) and (25) are sufficient, we may suppose that $\alpha = -\pi$. By Lemma 2, we can express (23) as the sum of an F_2A series in $(-\pi, \pi)$, plus an expression of the form

$$\gamma_0 + \gamma_1 e^{i\mu_1 x} + \gamma_{-1} e^{-i\mu_1 x}, \qquad (\gamma_{-1} = \bar{\gamma}_1), \tag{27}$$

where the γ's satisfy the conditions of the theorem, viz.

$$\sum_{k = \pm 1} \gamma_k e^{-i\pi\mu_k}(e^{2\pi i \mu_k} - 1) = 0, \tag{28}$$

and

$$\gamma_0/a + \sum_{k = \pm 1} \gamma_k e^{-i\pi\mu_k}(e^{2\pi i \mu_k} - 1)/(2ib\mu_k) = 0. \tag{29}$$

Now equation (28) is $(\gamma_1-\bar{\gamma}_1)2i\sin\pi\mu_1=0$, and $\sin\pi\mu_1\neq 0$ by (6). Thus γ_1 is real, and (29) can be written

$$\gamma_0/a+2\gamma_1\sin\pi\mu_1/b\mu_1=0. \tag{30}$$

Hence (27) is $2\gamma_1(-(a/b)\cdot\sin\pi\mu_1/\mu_1+\cos\mu_1 x)$, and, by Lemma 2, is an F_2A series in $(-\pi,\pi)$. This completes the proof.

5. The F_2A series of $f(x)$ in $(\alpha,\alpha+2\pi)$, being uniformly equiconvergent with the Fourier series of $f(x)$, is summable (C, 1) almost everywhere in the interval to $f(x)$. The first step in the investigation of the sum (C, 1) of the series outside the interval $(\alpha,\alpha+2\pi)$ is to determine the value of the sum (C, 1) in the interval $(\alpha+2\pi,\alpha+4\pi)$.

THEOREM 3. *If*

$$f(x)\sim\sum c_k e^{i\mu_k x}, \qquad (\alpha,\alpha+2\pi), \tag{31}$$

then $\sum c_k e^{i\mu_k x}(e^{2\pi i\mu_k}-1)$ *is uniformly convergent for* $\alpha\leqq x\leqq\alpha+2\pi$, *and its sum is*

$$-[2b/(m_1-m_2)][m_1e^{-m_1x}\int_\alpha^x f(t)e^{m_1t}\,dt-m_2e^{-m_2x}\int_\alpha^x f(t)e^{m_2t}\,dt]. \tag{32}$$

Proof. We multiply (31) by e^{m_1x} and integrate term by term from α to x. This is permissible for an F_2A series, since it is permissible for a Fourier series. We obtain

$$\int_\alpha^x f(t)e^{m_1t}\,dt=\sum c_k\int_\alpha^x e^{(i\mu_k+m_1)t}\,dt,$$

the series being uniformly convergent. Hence

$$\begin{aligned}\int_\alpha^x f(t)e^{m_1t}\,dt&=-i[\sum c_k e^{(i\mu_k+m_1)t}/(\mu_k-im_1)]_\alpha^x\\&=-i\sum c_k e^{(i\mu_k+m_1)x}/(\mu_k-im_1),\end{aligned}$$

in virtue of the first equation (16). Similarly,

$$\int_\alpha^x f(t)e^{m_2t}\,dt=-i\sum c_k e^{(i\mu_k+m_2)x}/(\mu_k-im_2).$$

Hence (32) equals

$$\begin{aligned}&[2ib/(m_1-m_2)]\sum c_k e^{i\mu_k x}[m_1/(\mu_k-im_1)-m_2/(\mu_k-im_2)]\\&\qquad=\sum c_k\, e^{i\mu_k x}2ib\mu_k/[(\mu_k-im_1)(\mu_k-im_2)]\\&=\sum c_k e^{i\mu_k x}(e^{2\pi i\mu_k}-1),\text{ by (13).}\end{aligned}$$

In the interval $(\alpha+2\pi, \alpha+4\pi)$, (31) is not necessarily an F_2A series. For in order that this shall be the case, it is necessary that the conditions obtained from (16) on replacing α by $\alpha+2\pi$ shall be satisfied. We shall presently see that by altering c_0 and c_1 in various ways, we can ensure that the new series is an F_2A series in $(\alpha+2\pi, \alpha+4\pi)$. It is the F_2A series of its sum (C, 1), which is known on account of Theorem 3 and the known changes in c_0 and c_1. Then, by Theorem 3 (applied in the interval $(\alpha+2\pi, \alpha+4\pi)$), we know its sum (C, 1) in $(\alpha+4\pi, \alpha+6\pi)$. Hence, we know the sum (C, 1) of the original series in $(\alpha+4\pi, \alpha+6\pi)$; and so on.

6. On account of the fact that the coefficients c_k of an F_2A series in $(\alpha, \alpha+2\pi)$ are connected by two equations, it is convenient to omit enough terms to ensure that the constants which remain are independent, and to direct our attention to the remaining series. It would be sufficient to omit terms with suffix $k, -k$ for any assigned $k \neq 0$. The simplest procedure however, seems to be to omit the three central terms with $k=-1, 0, 1$. We are thus led to the following definition.

DEFINITION. *A series*

$$\sum c_k e^{i\mu_k x}, \qquad c_{-k} = \bar{c}_k, \; |k| > 1, \tag{33}$$

is said to be an RF_2A series (reduced F_2A series) in $(\alpha, \alpha+2\pi)$, if there is a function $f(x) \varepsilon L(\alpha, \alpha+2\pi)$ such that $c_k = \lambda_k \int_\alpha^{\alpha+2\pi} f(t) e^{-i\mu_k t}\, dt$, $|k| > 1$.

We shall use the notation $\sum''$ to denote a sum in which the values $k=0, 1, -1$ are omitted. We shall indicate the relation between the RF_2A series in $(\alpha, \alpha+2\pi)$ and the function $f(x)$ by the notation

$$f(x) \approx \sum'' c_k e^{i\mu_k x}, \qquad (\alpha, \alpha+2\pi). \tag{34}$$

THEOREM 4. *If* (33) *is an RF_2A series in $(\alpha, \alpha+2\pi)$, then it is an RF_2A series in any interval $(\beta, \beta+2\pi)$.*

Proof. It suffices to consider the case in which $\alpha < \beta \leqq \alpha+2\pi$. (See the similar argument in [4], Theorem 1.) Given (34), we define c_r for $r=0, 1, -1$ by the equation

$$c_r = \lambda_r \int_\alpha^{\alpha+2\pi} f(t) e^{-i\mu_r t}\, dt.$$

Then, by Theorem 3, the series $\sum c_k e^{i\mu_k x}$ is summable (C, 1) almost everywhere in $(\alpha+2\pi, \alpha+4\pi)$ to

$$(35)\qquad \phi(x) = f(x-2\pi) - 2b/(m_1 - m_2)\cdot Q$$

where
$$Q = [m_1 e^{-m_1(x-2\pi)} \int_\alpha^{x-2\pi} f(t)\, e^{m_1 t}\, dt - m_2 e^{-m_2(x-2\pi)} \int_\alpha^{x-2\pi} f(t)\, e^{m_2 t}\, dt].$$

Hence, in the interval $(\beta, \beta+2\pi)$, the series $\sum c_k e^{i\mu_k x}$ is summable $(C, 1)$ almost everywhere to the function $F(x)$ defined by $F(x) = f(x)$ if $\beta \leqq x < \alpha + 2\pi$ and by $F(x) = \phi(x)$ if $\alpha + 2\pi \leqq x < \beta + 2\pi$. We shall now show that there is a complex number d_1 such that, if

$$(36)\qquad G(x) = F(x) + d_1 e^{i\mu_1 x} + d_{-1} e^{-i\mu_1 x}, \qquad (d_{-1} = \bar{d}_1),$$

then

$$(37)\qquad G(x) \approx {\sum}'' c_k e^{i\mu_k x}, \qquad (\beta, \beta+2\pi).$$

We first evaluate $I = \int_\beta^{\beta+2\pi} F(x)\, e^{-i\mu_k x}\, dx$. We have

$$(38)\qquad I = \int_\beta^{\alpha+2\pi} f(x)\, e^{-i\mu_k x}\, dx + \int_{\alpha+2\pi}^{\beta+2\pi} \phi(x)\, e^{-i\mu_k x}\, dx.$$

By (35), the last integral equals

$$e^{-2\pi i\mu_k}[\int_\alpha^\beta f(x) e^{-i\mu_k x}\, dx - (2bm_1/(m_1 - m_2)) \int_\alpha^\beta e^{-(m_1+i\mu_k)x}\, dx \int_\alpha^x f(t) e^{m_1 t}\, dt$$
$$+ (2bm_2/(m_1 - m_2)) \int_\alpha^\beta e^{-(m_2+i\mu_k)x}\, dx \int_\alpha^x f(t) e^{m_2 t}\, dt].$$

On changing the order of integration in the repeated integrals and using (13), the preceding expression becomes

$$(39)\qquad \int_\alpha^\beta f(t) e^{-i\mu_k t}\, dt + J_1(k) - J_2(k),$$

where

$$(40)\qquad J_r(k) = [2bm_r/(m_1 - m_2)\, e^{-2\pi i\mu_k} e^{-(m_r+i\mu_k)\beta}/(m_r + i\mu_k)] \int_\alpha^\beta f(t)\, e^{m_r t}\, dt.$$

To prove (37), we must show that

$$\int_\beta^{\beta+2\pi} G(x)\, e^{-i\mu_k x}\, dx = \int_\alpha^{\alpha+2\pi} f(x)\, e^{-i\mu_k x}\, dx, \qquad |k| > 1,$$

i.e., that

$$(41)\qquad \int_\beta^{\beta+2\pi} (d_1 e^{i\mu_1 x} + d_{-1} e^{-i\mu_1 x})\, e^{-i\mu_k x}\, dx + J_1(k) - J_2(k) = 0, \quad |k| > 1.$$

Now

$$(42)\qquad \int_\beta^{\beta+2\pi} e^{i(\mu_1-\mu_k)x}\, dx = e^{i(\mu_1-\mu_k)\beta}/[i(\mu_1 - \mu_k)]\, e^{-2\pi i\mu_k}(e^{2\pi i\mu_1} - e^{2\pi i\mu_k}).$$

On subtracting the equation

$$e^{2\pi i\mu_k} - 1 = [2b/(m_1 - m_2)][im_1/(\mu_k - im_1) - im_2/(\mu_k - im_2)]$$

from the similar equation with $k = 1$, and substituting in (42), we find that

$$\int_{\beta}^{\beta+2\pi} e^{i(\mu_1-\mu_k)x}\, dx = - e^{i(\mu_1-\mu_k)\beta} e^{-2\pi i\mu_k}[2b/(m_1 - m_2)]P,$$

where $P = m_1/(\mu_1 - im_1)(\mu_k - im_1) - m_2/(\mu_1 - im_2)(\mu_k - im_2)$. This equation holds when we replace μ_1 by $-\mu_1$. Substituting these results in (41), the condition to be satisfied takes the form

$$(43)\qquad [2b/(m_1 - m_2)]\, e^{-(2\pi+\beta)i\mu_k}[m_1N_1/(\mu_k - im_1) - m_2N_2/(\mu_k - im_2)] = 0,\qquad |k| > 1,$$

where, for $r = 1, 2$,

$$N_r = - d_1e^{i\mu_1\beta}/(\mu_1 - im_r) + d_{-1}e^{-i\mu_1\beta}/(\mu_1 + im_r) - ie^{-m_r\beta}\int_{\alpha}^{\beta} f(t)e^{m_rt}\, dt.$$

The condition (43) will therefore be satisfied, if we can choose d_1 so that

$$(44)\qquad N_1 = N_2 = 0.$$

If $\int_{\alpha}^{\beta} f(t)e^{m_rt}\, dt = 0$, $r = 1, 2$, we can choose $d_1 = 0$. In the contrary case, writing $z_r = d_1e^{i\mu_1\beta}/(\mu_1 - im_r)$, the equations (44) restrict z_r $(r = 1, 2)$ to lie on lines parallel to the real axis. Since $\arg(\mu_1 - im_1) \not\equiv \arg(\mu_1 - im_2)$, (mod. π), (44) is equivalent to the condition that d_1 shall lie on two non-parallel lines. This condition can therefore be satisfied.

7. In virtue of Theorem 4, we can use the expression "an RF_2A series" without specifying a particular interval of length 2π.

THEOREM 5. *Given an RF_2A series*

$$(45)\qquad \sum{}'' c_k e^{i\mu_kx}$$

and an interval $(\alpha, \alpha + 2\pi)$, the function $f(x)$ of which (45) is the RF_2A series in $(\alpha, \alpha + 2\pi)$ is determined to within an arbitrary additive multiple of

$$-(a/b)\cdot \sin \pi\mu_1/\mu_1 + \cos \mu_1 (x - \alpha - \pi).$$

Proof. We may suppose that $\alpha = -\pi$. By definition, there is a function $f_0(x)$ such that

$$(46)\qquad c_k = \lambda_k \int_{-\pi}^{\pi} f_0(x)e^{-i\mu_kx}\, dx,\qquad |k| > 1.$$

Since, by Lemma 2, $h(x) = -(a/b)\cdot\sin\pi\mu_1/\mu_1 + \cos\mu_1 x$ is its own F_2A series in $(-\pi,\pi)$, it follows that $c_k = \lambda_k\int_{-\pi}^{\pi}[f_0(x)+th(x)]e^{-i\mu_k x}\,dx$, $|k|>1$, for all t.

Conversely, let $f(x)$ be a function such that

$$c_k = \lambda_k\int_{-\pi}^{\pi} f(x)\,e^{-i\mu_k x}\,dx, \qquad (|k|>1). \tag{47}$$

By (46), $0 = \lambda_k\int_{-\pi}^{\pi}[f(x)-f_0(x)]e^{-i\mu_k x}\,dx$, $(|k|>1)$. Hence, the F_2A series of $f-f_0$ in $(-\pi,\pi)$ can be written as

$$\gamma_0 + \gamma_1 e^{i\mu_1 x} + \gamma_{-1}e^{-i\mu_1 x}, \qquad (\gamma_{-1}=\bar{\gamma}_1). \tag{48}$$

As in the proof of Theorem 2, γ_1 is real, say $\gamma_1 = \frac{1}{2}t$, and γ_0, γ_1 are connected by (30). Thus, (48) equals $th(x)$. This completes the proof.

Consider now a series $\sum c_k e^{i\mu_k x}$, $(c_{-k}=\bar{c}_k)$, given as the F_2A series of $f(x)$ in $(\alpha,\alpha+2\pi)$. What changes in c_0, c_1, c_{-1} are necessary and sufficient in order to obtain an F_2A series in $(\alpha+2\pi,\alpha+4\pi)$? This problem was referred to at the end of Section 5. By definition, (45) is an RF_2A series in $(\alpha,\alpha+2\pi)$. By Theorem 4, it is an RF_2A series in $(\alpha+2\pi,\alpha+4\pi)$. Hence there are numbers

$$\gamma_0, \gamma_1, \gamma_{-1}, \qquad (\gamma_{-1}=\bar{\gamma}_1), \tag{49}$$

such that

$$\gamma_0 + \gamma_1 e^{i\mu_1 x} + \gamma_{-1}e^{-i\mu_1 x} + \textstyle\sum'' c_k e^{i\mu_k x} \tag{50}$$

is an F_2A series in $(\alpha+2\pi,\alpha+4\pi)$. By Theorem 5, with α replaced by $\alpha+2\pi$, there is a unique triad of numbers (49) with $\gamma_0=0$; say $0,\gamma_1{}^*,\gamma_{-1}{}^*$. Write, for convenience, $\gamma_k{}^* = c_k$, $(|k|>1)$. Then $\gamma_1{}^*$ is uniquely determined by the appropriate form of equations (16); viz.

$$\sum \gamma_k{}^*\,e^{i\mu_k(\alpha+2\pi)}/(\mu_k - im_1) = 0, \qquad \sum \gamma_k{}^*\,e^{i\mu_k(\alpha+2\pi)}/(\mu_k - im_2) = 0.$$

It follows from Theorem 5, with α replaced by $\alpha+2\pi$, that the formulae for $\gamma_0, \gamma_1, \gamma_{-1}$ are

$$-t(a/b)\cdot\sin\pi\mu_1/\mu_1, \qquad \gamma_1{}^* + \tfrac{1}{2}t\,e^{-i\mu_1(\alpha+3\pi)}, \qquad \gamma_{-1}{}^* + \tfrac{1}{2}t\,e^{i\mu_1(\alpha+3\pi)}.$$

8. Another simple consequence of Theorem 5 is the following lemma, which will be used in the proof of the Young-Hausdorff inequalities for F_2A series.

LEMMA 3. *If* L, N, α *are real numbers, and* M *is a complex number such that*

$$(51)\qquad -L(a/b)\cdot\sin\pi\mu_1/\mu_1+\operatorname{Re}Me^{-i\mu_1(\alpha+\pi)}\neq 0,$$

then, given an RF_2A *series* (45), *there is precisely one triad* c_0, c_1, c_{-1} *with* c_0 *real,* $c_{-1}=\bar{c}_1$, *such that*

$$(52)\qquad Lc_0+Mc_1+\bar{M}c_{-1}=N,$$

and such that $\sum c_k e^{i\mu_k x}$ *is an* F_2A *series in* $(\alpha,\alpha+2\pi)$.

Proof. There are numbers (49) such that (50) is an F_2A series in $(\alpha,\alpha+2\pi)$. By Theorem 5, the general formulae for c_0, c_1, c_{-1}, such that $\sum c_k e^{i\mu_k x}$ is an F_2A series in $(\alpha,\alpha+2\pi)$, are

$$c_0=\gamma_0-t(a/b)\cdot\sin\ \pi\mu_1/\mu_1,\quad c_1=\gamma_1+\tfrac{1}{2}t\,e^{-i\mu_1(\alpha+\pi)},\quad c_{-1}=\gamma_{-1}+\tfrac{1}{2}t\,e^{i\mu_1(\alpha+\pi)}.$$

On substituting these values in (52), we obtain a linear equation in t. It has a unique solution, since the coefficient of t is the first member of (51).

9. The results of this section are proved in the same way as the analogous results concerning F. A. series.

THEOREM 6. *If* $f(x)\sim\sum c_k e^{i\mu_k x}$, $(\alpha,\alpha+2\pi)$, *and* $f\,\varepsilon\,L^2(\alpha,\alpha+2\pi)$, *then* $\int_\alpha^{\alpha+2\pi} f^2(x)\,dx=\sum|c_\nu|^2/\lambda_\nu$.

COROLLARY. *If* F *is a complex function, and* $F\,\varepsilon\,L^2(\alpha,\alpha+2\pi)$, *and*

$$d_\nu=\lambda_\nu\int_\alpha^{\alpha+2\pi}F(t)e^{-i\mu_\nu t}dt,\ \textit{then}\ \sum|d_\nu|^2/\lambda_\nu=\int_\alpha^{\alpha+2\pi}|F(t)|^2\,dt.$$

Proof. As for [4], Theorem 2, and Corollary.

LEMMA 4. *If* (45) *is an* RF_2A *series, then it is summable* (C, 1) *almost everywhere to a function* $g(x)$; *and, if* $v(x)$ *is a function of bounded variation in an assigned interval* (a, b), *then*

$$\lim_{n\to\infty}\int_\alpha^x v(t)\sum_{|\nu|<n}{}'' c_\nu e^{i\mu_\nu t}dt=\int_\alpha^x v(t)g(t)\,dt,$$

as $n\to\infty$, *uniformly for* $a\leqq x\leqq b$.

Proof. As for [4], Lemma 6.

THEOREM 7. *If* $g(x)$ *is the sum* (C, 1) *of the* RF_2A *series* (45), *then the series is the Fourier series of* $g(x)$; *i. e., if* $\mu\geqq 0$, *then*

$$\lim_{|T|\to\infty} 1/T \int_0^T g(x)e^{-i\mu x}dx = c_s \text{ if } \mu = \mu_s, s = 2, 3, \cdots,$$
$$= 0 \text{ otherwise.}$$

Proof. We need to know that if μ, $\geqq 0$, is assigned, then

$$\sum_{|k|>s} c_k/(\mu_k - \mu) \tag{53}$$

is convergent, s being a positive integer greater than 1. By Theorems 1 and 4, $\sum'' c_k e^{i\mu_k\alpha}/(\mu_k - im_1)$ is convergent for all α. In particular,

$$\sum_{|k|>s} c_k/(\mu_k - im_1) \tag{54}$$

is convergent. On subtracting (54) from (53), and using $c_k = o(1)$, we see that (53) is convergent. The proof proceeds as in [4], Theorem III. Note that here, $\mu_k = k - 1 + \eta_k$, while in [4], $\mu_k = k - \frac{1}{2} + \eta_k$, with $k\eta_k = O(1)$.

LEMMA 5. *If* $d_0, d_1, d_{-1}, \cdots, d_m, d_{-m}$ *are any complex numbers and* $1 < p \leqq 2$, $p^{-1} + q^{-1} = 1$, *then*

$$\left[\int_\alpha^{\alpha+2\pi} \Big|\sum_{|\nu|\leqq m} \lambda_\nu^{1/q} d_\nu e^{i\mu_\nu x}\Big|^q dx\right]^{1/q} \leqq \left[\sum_{|\nu|\leqq m} |d_\nu|^p\right]^{1/p}.$$

Proof. As for [4], Lemma 9.

10. We now consider the Young-Hausdorff inequalities for F_2A series. On account of the fact that the F_2A constants are connected by two equations, we cannot assign c_ν for all integer ν subject to $c_{-\nu} = \bar{c}_\nu$, $\sum |c_\nu|^p < \infty$, $1 < p \leqq 2$, and expect the c_ν to be the F_2A constants of a real function of class $L^q(\alpha, \alpha + 2\pi)$. On the other hand, if we omit c_0, c_1, c_{-1}, then there is one degree of freedom in the set of possible functions of which the c_ν are the F_2A constants. If, then, we require that the unspecified values of the three central constants shall satisfy an assigned linear equation, we may expect to find a unique function satisfying all the conditions.

THEOREM 8. (a) *Let* c_ν, $|\nu| > 1$ *be complex numbers such that* $\sum'' |c_\nu|^p < \infty$ *and* $c_{-\nu} = \bar{c}_\nu$. *Let* L, N, α *be assigned real numbers, and let* M *be an assigned complex number such that* (51) *holds. Then there is precisely one* (*real*) *function* $f \varepsilon L^q(\alpha, \alpha + 2\pi)$ *which satisfies the following two conditions*:

(i) $$c_\nu = \lambda_\nu \int_\alpha^{\alpha+2\pi} f(t) e^{-i\mu_\nu t} dt, \qquad |\nu| > 1;$$

(ii) *If* c_0, c_1, c_{-1} *are defined by the preceding equation for* $\nu = 0, 1, -1$, *then* $Lc_0 + Mc_1 + \bar{M}c_{-1} = N$. *For this function,*

$$(55) \qquad \left(\int_\alpha^{\alpha+2\pi} |f|^q dx\right)^{1/q} \leqq \left(\sum |c_\nu|^p/\lambda_\nu^{p-1}\right)^{1/p}.$$

(b) *If* $g(x)$ *is a real function of class* $L^p(\alpha, \alpha+2\pi)$, *and*

$$\gamma_\nu = \lambda_\nu \int_\alpha^{\alpha+2\pi} g(x)\, e^{-i\mu_\nu x} dx, \qquad (\nu = 0, \pm 1, \cdots),$$

then

$$\left(\sum |\gamma_\nu|^q/\lambda_\nu^{q-1}\right)^{1/q} \leqq \left(\int_\alpha^{\alpha+2\pi} |g|^p dx\right)^{1/p}.$$

Proof of (a). We define $c_0(n)$, $c_1(n)$ and $c_{-1}(n)$ by the equations

$$(56) \qquad Lc_0(n) + Mc_1(n) + \bar{M}c_{-1}(n) = N,$$

$$(57) \quad c_0(n)/(-im_1) + c_1(n)e^{i\mu_1\alpha}/(\mu_1 - im_1) + c_{-1}(n)e^{-i\mu_1\alpha}/(-\mu_1 - im_1) = -\sum_{|k|\leqq n}{}'' c_k e^{i\mu_k\alpha}/(\mu_k - im_1),$$

$$(58) \quad c_0(n)/(-im_2) + c_1(n)e^{i\mu_1\alpha}/(\mu_1 - im_2) + c_{-1}(n)e^{-i\mu_1\alpha}(-\mu_1 - im_2) = -\sum_{|k|\leqq n}{}'' c_k e^{i\mu_k\alpha}/(\mu_k - im_2).$$

The determinant of this system of equations is

$$(i \sin \pi\mu_1/ab)(m_1 - m_2)\{-L(2a/b)\cdot \sin \pi\mu_1/\mu_1 + Me^{-i\mu_1(\alpha+\pi)} + \bar{M}e^{i\mu_1(\alpha+\pi)}\} \neq 0.$$

The second members of (57), (58) are purely imaginary (or zero). Multiply (57) by im_1, (58) by im_2, and subtract. We obtain an equation of the form $c_1(n)A + c_{-1}(n)\bar{A} = B$, where B is real and $A \neq 0$. Hence $c_{-1}(n) = \bar{c}_1(n)$. Hence, in (57), all terms except the first are purely imaginary. Hence $c_0(n)$ is real. Since the series

$$\sum{}'' c_k e^{i\mu_k\alpha}/(\mu_k - im_1), \qquad \sum{}'' c_k e^{i\mu_k\alpha}(\mu_k - im_2),$$

converge with $\sum'' |c_\nu|^p$, it follows that there are numbers γ_0 (real), γ_1, γ_{-1} $(= \bar{\gamma}_1)$, such that $c_0(n) \to \gamma_0$, $c_1(n) \to \gamma_1$, $c_{-1}(n) \to \gamma_{-1}$, as $n \to \infty$. By (56),

$$(59) \qquad L\gamma_0 + M\gamma_1 + \bar{M}\gamma_{-1} = N.$$

Let $T_n(x) = c_0(n) + c_1(n)e^{i\mu_1 x} + c_{-1}(n)e^{-i\mu_1 x} + \sum_{|k|\leqq n}'' c_k e^{i\mu_k x}$. By Lemma 5, $\{T_n(x)\}$ converges strongly with index q to a real function, f say. By Theorem 2, and (57), (58), $T_n(x)$ is its own F_2A series in $(\alpha, \alpha + 2\pi)$. Hence, for a given integer ν with $|\nu| > 1$ and for $n > |\nu|$,

$$c_\nu/\lambda_\nu = \int_\alpha^{\alpha+2\pi} T_n(x)\,e^{-i\mu_\nu x}dx = \lim_{n\to\infty}\int_\alpha^{\alpha+2\pi} T_n(x)\,e^{-i\mu_\nu x}dx$$
$$= \int_\alpha^{\alpha+2\pi} f(x)\,e^{-i\mu_\nu x}dx.$$

Further, for $\nu = 0, 1, -1$,

$$\gamma_\nu = \lim_{n\to\infty} c_\nu(n) = \lim_{n\to\infty} \lambda_\nu \int_\alpha^{\alpha+2\pi} T_n(x)\,e^{-i\mu_\nu x}dx = \lambda_\nu \int_\alpha^{\alpha+2\pi} f(x)\,e^{-i\mu_\nu x}dx = c_\nu.$$

Thus, the conditions (i), (ii) are satisfied. Further, since, by Lemma 5,

$$\Big(\int_\alpha^{\alpha+2\pi} | T_n(x)|^q dx\Big)^{1/q} \leqq [| c_0(n)|^p/\lambda_0^{p-1} + | c_1(n)|^p/\lambda_1^{p-1}$$
$$+ | c_{-1}(n)|^p/\lambda_1^{p-1} + \sum_{|\nu|\leqq n}'' | c_\nu|^p/\lambda_\nu^{p-1}]^{1/p},$$

(55) follows on letting $n\to\infty$.

That there is only one function $f(x)$ which satisfies the conditions (i), (ii), follows from Lemma 3. For if f_0 is such a function, then, by Lemma 3, f_0 has the same F_2A constants as f. The F_2A series of $f_0 - f$ is thus identically zero. But $f_0 - f$ is, almost everywhere in $(\alpha, \alpha + 2\pi)$, the sum (C, 1) of its F_2A series. Hence $f_0 - f = 0$.

Proof of (b). As for [4], Theorem IV(b).

11. We now consider the simplest theorem of uniqueness for convergent series of the form $\sum c_k e^{i\mu_k x}$ with $c_{-k} = \bar{c}_k$. It is no loss of generality to suppose that the interval of convergence, given as of length 2π, is $(-\pi, \pi)$.

Consider the F_2A series of the function $f(x)$ defined as equal to 1 for $| x | \leqq \pi$. Then $1 \sim \sum \gamma_k e^{i\mu_k x}$, where $\gamma_0 = 2\pi\lambda_0$, $\gamma_k = 2\lambda_k \sin \pi\mu_k/\mu_k$, $(k \neq 0)$. By (1), $A_k^2/\cos^2 \pi\mu_k = B_k^2/\sin^2 \pi\mu_k = A_k^2 + B_k^2$. Hence

$$\mu_k^2 \sin^2 \pi\mu_k = b^2\mu_k^4/[(\mu_k^2 + a)^2 + b^2\mu_k^2] \to b^2, \tag{60}$$

as $| k | \to\infty$. It follows that

$$\gamma_0 - 1 + \sum' \gamma_k e^{i\mu_k x} \tag{61}$$

is a series whose coefficients are $O(k^{-2})$, and which converges to zero in the closed interval $(-\pi, \pi)$. We thus see that in order to obtain a significant uniqueness theorem, we must introduce some explicit restriction on the coefficients. It will be seen that the condition $c_0 = 0$ is such a restriction.

LEMMA 6. *If* $\sum' c_k e^{i\mu_k x}$, *where* $c_{-k} = \bar{c}_k$, $kc_k = o(1)$, *converges to zero almost everywhere in* $(-\pi, \pi)$, *then* $c_k = 0$ $(k = \pm 1, \pm 2, \cdots)$.

Proof. By hypothesis, $\sum'' c_k e^{i\mu_k x}$ converges almost everywhere in $(-\pi, \pi)$ to $-c_1 e^{i\mu_1 x} - c_{-1} e^{-i\mu_1 x}$. In Theorem 8, let $p = 2$, $M = N = 0$, $\alpha = -\pi$, $L = 1$. By the first part of that theorem, there are numbers c_1^*, c_{-1}^* $(= \bar{c}_1^*)$ such that

$$c_1^* e^{i\mu_1 x} + c_{-1}^* e^{-i\mu_1 x} + \sum'' c_k e^{i\mu_k x} \tag{62}$$

is an F_2A series in $(-\pi, \pi)$. But this series converges almost everywhere in $(-\pi, \pi)$ to

$$(c_1^* - c_1) e^{i\mu_1 x} + (c_{-1}^* - c_{-1}) e^{-i\mu_1 x}. \tag{63}$$

Hence (62) is the F_2A series of (63) in $(-\pi, \pi)$. Write

$$c_1^* - c_1 = \tfrac{1}{2}(\alpha - i\beta). \tag{64}$$

Then (63) is $\alpha \cos \mu_1 x + \beta \sin \mu_1 x$. Since the constant term in (62) is zero,

$$0 = \lambda_0 \int_{-\pi}^{\pi} (\alpha \cos \mu_1 x + \beta \sin \mu_1 x)\, dx.$$

Hence, $\alpha = 0$. For $|k| > 1$, we have

$$c_k/\lambda_k = \int_{-\pi}^{\pi} \beta \sin \mu_1 x\, e^{-i\mu_k x}\, dx = 2i\beta/b \cdot \sin \pi\mu_1 \sin \pi\mu_k,$$

by (20). By (7) and (60), $|\mu_k c_k| \to \frac{1}{\pi} |\beta \sin \pi\mu_1|$. By hypothesis, $\mu_k c_k \to 0$. Further, by (6), μ_1 is not an integer. Hence $\beta = 0$. Thus, all the coefficients in (62) are zero. By (64), $c_1 = c_1^*$. Hence all the c_k are zero.

The series obtained by differentiating (61) term by term, viz.

$$\sum' i\mu_k \gamma_k\, e^{i\mu_k x}, \tag{65}$$

has its coefficients $O(k^{-1})$. As in the proof of Lemma 6, if we alter the two central terms suitably, we obtain an F_2A series, say

$$\phi(x) \sim i\mu_1 \gamma_1^* e^{i\mu_1 x} - i\mu_1 \gamma_{-1}^* e^{-i\mu_1 x} + \sum'' i\mu_k \gamma_k e^{i\mu_k x}, \qquad (-\pi, \pi). \tag{66}$$

This series can be integrated term by term from 0 to x. Since (61) converges to zero, it follows that

$$\int_0^x \phi(u)\, du - (\gamma_1^* - \gamma_1) e^{i\mu_1 x} - (\gamma_{-1}^* - \gamma_{-1}) e^{-i\mu_1 x}$$

is constant in $|x| \leqq \pi$. Hence we may take $\phi(x)$ to be

$$i\mu_1 (\gamma_1^* - \gamma_1) e^{i\mu_1 x} - i\mu_1 (\gamma_{-1}^* - \gamma_{-1}) e^{-i\mu_1 x},$$

for $|x| \leqq \pi$. On account of the equiconvergence of an F_2A series in $(-\pi, \pi)$ with a Fourier series in $(-\pi, \pi)$, the second member of (66) converges to this function, and so (65) converges to zero for $|x| \leqq \pi$. That it is essentially the only series without a constant term which has this property, is shown by the following theorem.

THEOREM 9. *If* $\sum' c_k e^{i\mu_k x}$, *where* $c_{-k} = \bar{c}_k$, *converges to zero for* $|x| < \pi$, *then there is a real constant* c *such that* $c_k = ci\mu_k\gamma_k$.

Proof. The hypothesis of convergence implies that $c_k = o(1)$. Let

$$F(x) = -\sum c_k/\mu_k^2 \cdot e^{i\mu_k x}. \tag{67}$$

By the classical argument of Riemann, $D^2F(x) = 0$ for $|x| < \pi$, where $D^2F(x)$ denotes the generalised second derivative. Hence, $F(x) = lx + m$, where l, m are real numbers.

With γ_k defined as above, write $d_k = \gamma_k/(1-\gamma_0)$, $(k \neq 0)$. Then

$$\sum' d_k e^{i\mu_k x} = 1, \qquad (|x| \leqq \pi), \tag{68}$$

and $|\mu_k^2 d_k| = |2\lambda_k\mu_k \sin \pi\mu_k/(1-\gamma_0)| \to |b/\pi(1-\gamma_0)|$, as $|k| \to \infty$. The series (68) is uniformly convergent. Hence, $\sum' d_k/i\mu_k \cdot e^{i\mu_k x} = x$ $(|x| \leqq \pi)$, and

$$\sum' d_k(l/(i\mu_k) + m)e^{i\mu_k x} = lx + m. \tag{69}$$

The difference between the series in (69) and (67) converges to zero for $|x| < \pi$, and has coefficients which are $O(k^{-2})$. By Lemma 6, they are all zero. Thus $-c_k/\mu_k^2 = d_k(m + l/i\mu_k)$. Multiply by μ_k^2 and let $k \to \infty$. Then $0 = |mb/[\pi(1-\gamma_0)]|$. Thus, $m = 0$, and $c_k = ild_k\mu_k = l/(1-\gamma_0) \cdot i\mu_k\gamma_k$, as required.

Sections 12-16. F_2B Series.

12. The theory of F_2B series can be developed on lines similar to the preceding theory of F_2A series. The details of the analysis, however, are sufficiently different to justify the short sketch which follows. The notation will be almost the same as for F_2A series, but the meanings of the symbols will be different.

The F_2B series (see Section 1) is a series of cosines and sines of $\mu_k x$ where the μ_k are the roots, supposed all real and simple, of the equation

$$A \sin \pi z - B \cos \pi z = 0, \tag{70}$$

where $A = az$, $B = z^2 + b$, $ab \neq 0$. Then $z = 0$ is not a root of (70).

If $f(x)$ is a real function of class $L(\alpha, \alpha + 2\pi)$, then the F_2B series of $f(x)$ in $(\alpha, \alpha + 2\pi)$ can be written as

$$f(x) \sim \sum' c_k e^{i\mu_k x}, \qquad (\alpha, \alpha + 2\pi), \tag{71}$$

where $c_k = \lambda_k \int_\alpha^{\alpha+2\pi} f(x) e^{-i\mu_k x} dx$, $(k = \pm 1, \pm 2, \cdots)$, and $\mu_1 < \mu_2 < \cdots$ are the positive roots of the equation (70), $\mu_{-k} = -\mu_k$, $c_{-k} = \bar{c}_k$, and λ_k is the residue at $z = \mu_k$ of

$$Q(z) = \tfrac{1}{2}(A \cos \pi z + B \sin \pi z)/(A \sin \pi z - B \cos \pi z).$$

We find that

$$(72) \qquad \mu_k = k - \tfrac{3}{2} + \eta_k, \qquad k\eta_k = O(1).$$

Writing $A_k = A(\mu_k)$, etc., we have

$$1/(2\lambda_k) = \pi + (A_k' \sin \pi\mu_k - B_k' \cos \pi\mu_k)/(A_k \cos \pi\mu_k + B_k \sin \pi\mu_k).$$

But $A_k/\cos \pi\mu_k = B_k/\sin \pi\mu_k = A_k \cos \pi\mu_k + B_k \sin \pi\mu_k$. Hence

$$(73) \qquad 1/(2\lambda_k) = \pi + (a \sin \pi\mu_k - 2\mu_k \cos \pi\mu_k) \cos \pi\mu_k / a\mu_k.$$

If ν, n are positive integers, we have

$$(74) \qquad \int_{-\pi}^{\pi} \cos \mu_\nu x \cos \mu_n x \, dx = (2/a) \cdot \cos \pi\mu_\nu \cos \pi\mu_n, \qquad (n \neq \nu),$$

$$(75) \qquad \int_{-\pi}^{\pi} \sin \mu_\nu x \sin \mu_n x \, dx = (-2b/a) \cdot (\cos \pi\mu_\nu/\mu_\nu) \cdot (\cos \pi\mu_n/\mu_n), \qquad (n \neq \nu).$$

13. The F_2B series (71) is uniformly equiconvergent in the closed interval $(\alpha, \alpha + 2\pi)$ with a Fourier series, though not with the Fourier series of a function of period 2π. We have the following special case of a known theorem ([3], Theorem 1).

LEMMA 7. *If $s_n(x)$ is the n-th partial sum of (71), i. e., $s_n = \sum_{|k| \leqq n}' c_k e^{i\mu_k x}$, and if $\tau_{2n}(x)$ is the 2n-th partial sum of the Fourier series of the function $\phi(x)$, of period 4π, defined by $\phi(x) = f(x)$ if $\alpha \leqq x \leqq \alpha + 2\pi$ and by $\phi(x) = -f(x - 2\pi)$ if $\alpha + 2\pi < x < \alpha + 4\pi$, then $s_n(x) - \tau_{2n}(x) \to 0$ uniformly in $(\alpha, \alpha + 2\pi)$ as $n \to \infty$. In particular, $s_n(\alpha + 2\pi) + s_n(\alpha) \to 0$ as $n \to \infty$.*

According to the last clause of Lemma 7,

$$(76) \qquad \sum' c_k e^{i\mu_k x}(e^{2\pi i\mu_k} + 1) = 0.$$

Equation (70) can be written $e^{2\pi i z} = (A + iB)/(A - iB)$. Hence

$$(77) \qquad e^{2\pi i\mu_k} + 1 = 2ia\mu_k/(\mu_k^2 + ia\mu_k + b) = 2ia\mu_k/[(\mu_k + im_1)(\mu_k + im_2)].$$

We shall suppose that $a^2 + 4b > 0$, so that m_1, m_2 are real and different. We shall now prove the analogue of Theorem 1.

THEOREM 10. *For any F_2B series* (71), *we have*

$$\sum{}' c_k e^{i\mu_k\alpha}\mu_k/[(\mu_k+im_1)(\mu_k+im_2)]=0, \tag{78}$$

and

$$\sum{}' c_k e^{i\mu_k\alpha}/[(\mu_k+im_1)(\mu_k+im_2)]=0, \tag{79}$$

or, what is equivalent,

$$\sum{}' c_k e^{i\mu_k\alpha}/(\mu_k+im_1)=0, \qquad \sum{}' c_k e^{i\mu_k\alpha}/(\mu_k+im_2)=0. \tag{80}$$

Proof. Equation (78) follows from (76) on substituting the value of $e^{2\pi i\mu_k}+1$ from (77). To prove (79), or, what is equivalent,

$$\sum{}' c_k/\mu_k\cdot e^{i\mu_k\alpha}(e^{2\pi i\mu_k}+1)=0, \tag{81}$$

it suffices to show that

$$\sum{}' \lambda_k/\mu_k\cdot\cos\pi\mu_k e^{i\mu_k(x-\alpha-\pi)}=0, \tag{82}$$

uniformly in $(\alpha,\alpha+2\pi)$. For then, on taking conjugates, multiplying by $f(x)$ and integrating over $(\alpha,\alpha+2\pi)$, we obtain (81).

Changing the variable to $x-\alpha-\pi$, we have to prove that

$$\sum_1^\infty \lambda_n/\mu_n\cdot\cos\pi\mu_n\sin\mu_n x=0, \tag{83}$$

uniformly in $(-\pi,\pi)$.

Consider the F_2B series, in the interval $(-\pi,\pi)$, of the function $\sin\mu_1 x$; say

$$\sum_1^\infty b_k\sin\mu_k x. \tag{84}$$

For $k>1$, we have by (75),

$$b_k=-4b\lambda_k/a\cdot\cos\pi\mu_1/\mu_1\cdot\cos\pi\mu_k/\mu_k. \tag{85}$$

Further,

$$b_1=2\lambda_1(\pi-1/\mu_1\cdot\sin\pi\mu_1\cos\pi\mu_1),$$

$$\begin{aligned} b_1-1&=2\lambda_1(\pi-1/(2\lambda_1)-1/\mu_1\cdot\sin\pi\mu_1\cos\pi\mu_1)\\ &=2\lambda_1(-2\cos\pi\mu_1/\mu_1)(\sin\pi\mu_1-\mu_1/a\cdot\cos\pi\mu_1)\\ &=-4b\lambda_1/a(\cos\pi\mu_1/\mu_1)^2, \end{aligned} \tag{86}$$

by (73) and (70). By Lemma 7, the series (84) converges uniformly to $\sin\mu_1 x$. Substituting the values (85), (86), we obtain (83). Clearly, (78), (79), are equivalent to (80).

14. LEMMA 8. *If ν is a positive integer > 1, then*

$$f_\nu(x) = -\cos\mu_1 x/\cos\mu_1\pi + \cos\mu_\nu x/\cos\mu_\nu\pi,$$

is its own F_2B series in $(-\pi, \pi)$; and

$$g_\nu(x) = -\mu_1 \sin\mu_1 x/\cos\mu_1\pi + \mu_\nu \sin\mu_\nu x/\cos\mu_\nu\pi,$$

is its own F_2B series in $(-\pi, \pi)$.

Proof. The F_2B series of $f_\nu(x)$ can be written $\sum_1^\infty a_n\cos\mu_n x$. If $n \neq 1, \nu$, then $a_n = 2\lambda_n \int_{-\pi}^{\pi} f_\nu(x)\cos\mu_n x\,dx = 0$, by (74). Hence the F_2B series of $f_\nu(x)$ is $a_1\cos\mu_1 x + a_\nu\cos\mu_\nu x$. This converges to $f_\nu(x)$ for $|x| < \pi$. Giving x the value $\pi/2\mu_\nu$, we find $a_1 = -\sec\mu_1\pi$. Hence $a_\nu = \sec\mu_\nu\pi$. The proof of the second part is similar.

THEOREM 11. *In order that*

$$(87) \qquad \sum_{|k|\leqq n}{}' c_k e^{i\mu_k x}, \qquad c_{-k} = \bar{c}_k,$$

shall be an F_2B series in $(\alpha, \alpha + 2\pi)$, it is necessary and sufficient that

$$(88) \qquad \sum_{|k|\leqq n}{}' c_k e^{i\mu_k\alpha}(e^{2\pi i\mu_k} + 1) = 0,$$

and that

$$(89) \qquad \sum_{|k|\leqq n}{}' c_k/\mu_k \cdot e^{i\mu_k\alpha}(e^{2\pi i\mu_k} + 1) = 0.$$

Proof. By Theorem 10 and (77), the conditions (88), (89) are necessary. To prove that they are sufficient, we may suppose that $\alpha = -\pi$. Then the conditions become, on using the notation $c_k = \frac{1}{2}(a_k - ib_k)$,

$$\sum_1^n a_k\cos\pi\mu_k = 0, \qquad \sum_1^n b_k/\mu_k \cdot \cos\pi\mu_k = 0.$$

Hence $\sum_1^n a_k\cos\mu_k x + b_k\sin\mu_k x$ becomes

$$\sum_2^n a_k(\cos\mu_k x - \cos\mu_k\pi/(\cos\mu_1\pi)\cdot\cos\mu_1 x)$$
$$+ \sum_2^n b_k(\sin\mu_k x - \mu_1\cos\mu_k\pi/(\mu_k\cos\mu_1\pi)\cdot\sin\mu_1 x).$$

By Lemma 8, each term is an F_2B series in $(-\pi, \pi)$. So then is the first member, i.e. (87).

Enough has now been said to make it clear that F_2B series may be investigated in the same way as F_2A series. Some results are simpler, since, in an F_2B series, c_1, c_{-1} are determined by the remaining coefficients, and there is no constant term. We shall not enunciate the analogues of Theorems 3-8. We shall merely enunciate a special case of the analogue of Theorem 8(a), which will be used below.

LEMMA 9. *Let* c_ν, $|\nu| > 1$, *be complex numbers such that* $\sum'' |c_\nu|^2 < \infty$, $c_{-\nu} = \bar{c}_\nu$. *Given the interval* $(\alpha, \alpha + 2\pi)$, *there is precisely one real function* f *of class* $L^2(\alpha, \alpha + 2\pi)$, *such that* $c_\nu = \lambda_\nu \int_\alpha^{\alpha+2\pi} f(t) e^{-i\mu_\nu t} dt$, $(|\nu| > 1)$.

15. We now turn to the problem of the uniqueness of convergent series of the form $\sum' c_k{}^{i\mu_k x}$ with $c_{-k} = \bar{c}_k$. We shall prove (Theorem 12) that any series of this form which converges to zero in the open interval $(-\pi, \pi)$, is a linear combination of two particular series with this property. The particular series are (83), and the series obtained from it by termwise differentiation. We shall use the fact that, since $\tan \pi\mu_k = (\mu_k{}^2 + b)/a\mu_k$, we have, by (72),

$$\tan \pi\eta_k = -a\mu_k/(\mu_k{}^2 + b) = -a/\mu_k + O(k^{-3}),$$

$$\eta_k = -a/\pi\mu_k + O(k^{-3}),$$

$$\cos \pi\mu_k = (-1)^{k-1} \sin \pi\eta_k = (-1)^k a/\mu_k + O(k^{-3}),$$

and

$$(90) \qquad |\mu_k \cos \pi\mu_k| \to |a|, \qquad \text{as } k \to \infty.$$

16. LEMMA 10. *If*

$$(91) \qquad \sum' c_k e^{i\mu_k x}, \qquad c_{-k} = \bar{c}_k, \, kc_k = o(1),$$

converges uniformly to zero in $(-\pi, \pi)$, *then there is a real number* l *such that*

$$(92) \qquad c_k = l(i\lambda_k/\mu_k) \cdot \cos \pi\mu_k;$$

i.e., (91) *is a constant multiple of* (83).

Proof. The series $\sum'' c_k e^{i\mu_k x}$ converges uniformly to $-c_1 e^{i\mu_1 x} - c_{-1} e^{-i\mu_1 x}$. By Lemma 9, there are numbers $c_1{}^*$, $c_{-1}{}^*$ $(= \bar{c}_1{}^*)$ such that

$$c_1{}^* e^{i\mu_1 x} + c_{-1}{}^* e^{-i\mu_1 x} + \sum'' c_k e^{i\mu_k x}$$

is an F_2B series in $(-\pi, \pi)$. It is the F_2B series of its sum, viz.,

$$(c_1{}^* - c_1) e^{i\mu_1 x} + (c_{-1}{}^* - c_{-1}) e^{-i\mu_1 x}.$$

Write $c_1{}^* - c_1 = \frac{1}{2}(\alpha - i\beta)$. Then, for $|k| > 1$,

$$
\begin{aligned}
(93) \qquad c_k/\lambda_k &= \int_{-\pi}^{\pi} (\alpha \cos \mu_1 x + \beta \sin \mu_1 x) e^{-i\mu_k x}\, dx \\
&= \alpha \int_{-\pi}^{\pi} \cos \mu_1 x \cos \mu_k x\, dx - i\beta \int_{-\pi}^{\pi} \sin \mu_1 x \sin \mu_k x\, dx \\
&= 2\alpha/a \cdot \cos \mu_1 \pi \cos \mu_k \pi + 2ib\beta/a \cdot \cos \mu_1 \pi/\mu_1 \cdot \cos \mu_k \pi/\mu_k,
\end{aligned}
$$

by (74), (75). By hypothesis, $kc_k \to 0$. By (70), $\cos \pi\mu_1 \neq 0$. Multiply (93) by μ_k, and let $k \to \infty$. By (90), $\alpha = 0$. Hence we have (92) for $|k| > 1$, with $l = 2b\beta/a \cdot \cos \mu_1 \pi/\mu_1$. Since $c_1 = \frac{1}{2}(a_1 - ib_1)$, (91) is

$$
a_1 \cos \mu_1 x + b_1 \sin \mu_1 x - 2l \sum_{2}^{\infty} \lambda_n/\mu_n \cdot \cos \mu_n \pi \sin \mu_n x,
$$

and converges to $a_1 \cos \mu_1 x + b_1 \sin \mu_1 x + 2l(\lambda_1/\mu_1) \cdot \cos \mu_1 \pi \sin \mu_1 x$, by (83). By hypothesis, this is zero. Hence $a_1 = 0$, $b_1 = -2l(\lambda_1/\mu_1) \cdot \cos \mu_1 \pi$, as required.

LEMMA 11. *The series*

$$
(94) \qquad \sum{}' \lambda_k \cos \pi\mu_k\, e^{i\mu_k x}
$$

converges boundedly to zero in the open interval $(-\pi, \pi)$.

Proof. By Lemma 9, there are numbers $\lambda_1{}^*$, $\lambda_{-1}{}^*$ $(= \bar{\lambda}_1{}^*)$ such that

$$
(95) \qquad \lambda_1{}^* e^{i\mu_1 x} + \lambda_{-1}{}^* e^{-i\mu_1 x} + \sum{}'' \lambda_k \cos \pi\mu_k e^{i\mu_k x}
$$

is an F_2B series in $(-\pi, \pi)$. Hence it is summable $(C, 1)$ almost everywhere in that interval, and can be integrated term by term. The same is therefore true of (94). Denote the sum $(C, 1)$ of (94) by $f(x)$. Then

$$
\sum{}' \lambda_k/(i\mu_k) \cdot \cos \pi\mu_k e^{i\mu_k x} = \int_0^x f(t)\, dt.
$$

By (83) the first member is zero. Hence $f = 0$ almost everywhere. Then (95) is the F_2B series of

$$
(96) \qquad (\lambda_1{}^* - \lambda_1) e^{i\mu_1 x} + (\lambda_{-1}{}^* - \lambda_{-1}) e^{-i\mu_1 x}.
$$

By Lemma 7, (95) converges boundedly in the open interval $(-\pi, \pi)$ to (96). Hence (94) converges boundedly in this interval to zero.

We finally have,

THEOREM 12. *If* $\sum{}' c_k e^{i\mu_k x}$, *where* $c_{-k} = \bar{c}_k$, *converges to zero in the open interval* $(-\pi, \pi)$, *then it is a linear combination of the two series*

$$\sum{}' \lambda_k \cos \pi\mu_k e^{i\mu_k x}, \qquad \sum{}' \lambda_k/(i\mu_k) \cdot \cos \pi\mu_k e^{i\mu_k x},$$

of which the first converges boundedly, and the second converges uniformly to zero in the open interval.

Proof. As for Theorem 9, $-\sum{}' c_k/\mu_k^2 \cdot e^{i\mu_k x} = lx + m$ with real constant l and m. On integrating (83), we infer that $\sum{}' \lambda_k/\mu_k^2 \cdot \cos \pi\mu_k e^{i\mu_k x}$ converges uniformly to a constant p. By Lemma 10, this constant is not zero. Then $\sum{}' \lambda_k/(i\mu_k^3) \cdot \cos \pi\mu_k e^{i\mu_k x}$ converges uniformly to px. Hence

$$\sum [l/p \cdot \lambda_k/(i\mu_k^3) \cdot \cos \pi\mu_k + m/p \cdot \lambda_k/\mu_k^2 \cos \pi\mu_k + c_k/\mu_k^2] e^{i\mu_k x}$$

converges uniformly to zero. The coefficients are $o(k^{-2})$. By Lemma 10, there is a real number t such that

$$c_k/\mu_k^2 + m/p \cdot \lambda_k/\mu_k^2 \cos \pi\mu_k + l/p \cdot \lambda_k/(i\mu_k^3) \cdot \cos \pi\mu_k = ti(\lambda_k/\mu_k) \cdot \cos \pi\mu_k.$$

Since $c_k = o(1)$, it follows from (90) that $t = 0$. This proves the theorem.

Appendix.

17. The simple arguments of this paper, depend on the fact that the integrals

$$\int_{-\pi}^{\pi} \cos \mu_\nu x \cos \mu_n x \, dx, \qquad \int_{-\pi}^{\pi} \sin \mu_\nu x \sin \mu_n x \, dx,$$

where ν, n are different positive integers, can be expressed as a product of a function of μ_ν and a function of μ_n by means of the equation (1) which μ_ν and μ_n satisfy. The question arises: For what relatively prime polynomials A, B with real coefficients, one of which is even and the other odd, have the above integrals the form described?

For $\nu \neq n$ we have

$$\begin{aligned}\int_{-\pi}^{\pi} \cos \mu_\nu x \cos \mu_n x \, dx &= \sin \mu_\nu\pi \cos \mu_n\pi \, (2\mu_\nu/\mu_\nu^2 - \mu_n^2) \\ &\quad + \sin \mu_n\pi \cos \mu_\nu\pi \, (2\mu_n/\mu_n^2 - \mu_\nu^2) \\ &= 2 \cos \mu_\nu\pi \cos \mu_n\pi/(\mu_\nu^2 - \mu_n^2) \, [\mu_\nu B(\mu_\nu)/A(\mu_\nu) - \mu_n B(\mu_n)/A(\mu_n)],\end{aligned}$$

by (1). We therefore require that the value of the symmetric polynomial

$$P(x, y) = (yB(y)A(x) - xB(x)A(y))/(x^2 - y^2) \tag{97}$$

shall be of the form $b_\nu b_n$ when $x = \mu_\nu$, $y = \mu_n$, $n \neq \nu$. We shall prove (see Lemma 11 below), that this can only happen if $P(x, y) = f(x) f(y)$, where $f(x)$ is a polynomial.

Again,

$$\int_{-\pi}^{\pi} \sin \mu_\nu x \sin \mu_n x \, dx$$
$$= 2 \cos \mu_\nu \pi \cos \mu_n \pi / (\mu_\nu{}^2 - \mu_n{}^2) \, [\mu_n B(\mu_\nu)/A(\mu_\nu) - \mu_\nu B(\mu_n)/A(\mu_n)].$$

We therefore require that the value of the symmetric polynomial

$$(98) \qquad Q(x,y) = (xA(x)B(y) - yA(y)B(x))/(x^2 - y^2)$$

shall be of the form $c_\nu c_n$ for $x = \mu_\nu$, $y = \mu_n$, $\nu \neq n$. According to Lemma 11, this can only happen if $Q(x,y) = g(x)g(y)$, where $g(x)$ is a polynomial.

There are two possibilities. (i) A even, B odd. Put $x = 0$ in P. Then $A(0)B(y)/y = -f(0)f(y)$. Now $A(0) \neq 0$, since $(A,B) = 1$. Then $f(0) \neq 0$. Hence

$$(99) \qquad yB(y)A(x) - xB(x)A(y) = (y^2 - x^2)f(x)/f(0) \cdot B(y)/y \cdot A(0).$$

Hence $B(y)/y$ divides $A(y)$. Since $(A,B) = 1$, $B(y) = by$ with a constant b. By (99), A is a quadratic. Suppose that its leading coefficient is 1. Then there are two cases, $A(y) = y^2 + a$, $A(y) = 1$.

Consider next, the possibility (ii), A odd, B even. We put $x = 0$ in Q and find that $A(y) = ay$ with a constant a, and that B is a quadratic. If its leading coefficient is 1, then there are two cases, $B(y) = y^2 + b$, $B(y) = 1$.

Thus, the only cases in which the integrals are of the required form, are the two cases considered in this paper, and the two cases for which (1) becomes

$$(\alpha) \qquad \sin \pi z - bz \cos \pi z = 0, \qquad\qquad (\beta) \qquad az \sin \pi z - \cos \pi z = 0,$$

respectively. The case (α) has been the subject of recent investigations (see the remarks in [4], 324). The case (β) can be treated on similar lines.

We now turn to the lemma mentioned above. We note first, that when A is even and B is odd, $P(x,y)$ and $Q(x,y)/(xy)$ are symmetric polynomials in ξ $(= x^2)$ and η $(= y^2)$; and that when A is odd and B is even, the same is true of $P(x,y)/(xy)$ and $Q(x,y)$. It therefore suffices to prove

Lemma 11. *If $p(\xi,\eta)$ is a symmetric polynomial, not identically zero, and*

$$(100) \qquad p(a_r, a_s) = b_r b_s, \qquad (r \neq s),$$

where $0 < a_1 < a_2 < \cdots$, then $p(\xi,\eta) = c(\xi)c(\eta)$, where $c(\xi)$ is a polynomial in ξ.

Proof. If, for some r, $b_r = 0$, then $p(a_r, a_s) = 0$, $(s \neq r)$. Hence, $p(a_r, \eta) \equiv 0$. Let $(\xi - a_r)^n$ be the highest power of $(\xi - a_r)$ which divides $p(\xi, \eta)$. We consider $p(\xi, \eta)[(\xi - a_r)(\eta - a_r)]^{-n}$, and omit a_r from the sequence $a_1, a_2, \cdots$. Thus, we may suppose that $b_r \neq 0$ $(r = 1, 2, \cdots)$.

Then, if $r \neq s$, $p(a_t, a_r)/b_r = p(a_t, a_s)/b_s$ for $t \neq r$, $t \neq s$. Hence

$$p(\xi, a_r)/b_r = p(\xi, a_s)/b_s \tag{101}$$

for all ξ. Write

$$c(\xi) = p(\xi, a_1)/b_1. \tag{102}$$

Then by (101), on putting $s = 1$,

$$p(\xi, a_r) = b_r c(\xi) \tag{103}$$

for $r = 2, 3, \cdots$, and also for $r = 1$, by (102).

For a given s, choose an $r \neq s$ in (103). Put $\xi = a_s$. By (100), $c(a_s) = b_s$ for $s = 1, 2, \cdots$. Thus, by (103),

$$p(\xi, a_r) = c(a_r) c(\xi), \qquad (r = 1, 2, \cdots), \tag{104}$$

and so $p(\xi, \eta) = c(\xi) c(\eta)$.

18. *Historical note.* The investigations of this paper are based on [3], Theorem 1. There are points of similarity between the proof of that theorem and the investigations of Dini [1], who, writing before the era of the Lebesgue integral, could not easily attain to general and precise results. In a subsequent paper [2], Dini considered the problem of deciding when the equation (1) has all its roots real and simple. He examined the equations which form the subject of [3], Theorems 2, 3, 4; but his results are fragmentary.

19. (*Added 1 May, 1955*).

The series we have considered in this paper, can be written in the form

$$\int_{-\infty}^{\infty} e^{ixu}\, d\phi(u) = \lim_{\omega \to \infty} \int_{-\omega}^{\omega} e^{ixu} d\phi(u), \tag{105}$$

where $\phi(u)$ is constant in each open interval (μ_k, μ_{k+1}), $k = 0, \pm 1, \pm 2, \cdots$, and $\phi(\mu_k +) - \phi(\mu_k -) = c_k$. A theory of integrals (105) has been developed ([5], [7]) partly with the object of enabling one to deduce analogues of theorems concerning ordinary trigonometric series

$$\sum c_n e^{inx} \qquad c_{-n} = \bar{c}_n \tag{106}$$

without having to write out analogous proofs. As pointed out by the referee, this theory has a direct bearing on some of our results, in particular, Theorems 9 and 12.

Consider for example, the following proposition which generalizes Theorem 9: *If*

$$(107) \qquad \sum{}' c_k e^{i\mu_k x} \qquad c_{-k} = \bar{c}_k$$

converges to $f(x)$ for $|x| < \pi$, where $f(x)$ is finite and $\varepsilon L(-\pi, \pi)$, then there is a real constant c such that

$$c_k = ci\mu_k\gamma_k + g_0\gamma_k/(1-\gamma_0) + g_k, \qquad (k = \pm 1, \pm 2, \cdots),$$

where $g_k = \lambda_k \int_{-\pi}^{\pi} f(t) e^{-i\mu_k t}\, dt$, $(k = 0, \pm 1, \pm 2, \cdots)$.

To prove this, we observe that the hypothesis implies that $c_k = o(1)$. Let ϵ be a small positive number. By [7], Theorem 4, there is a series

$$(108) \qquad \sum \omega_n e^{inx} \qquad \omega_{-n} = \bar{\omega}_n,\ \omega_n = o(1),$$

such that (107), (108) are uniformly equiconvergent in $(-\pi + \epsilon, \pi - \epsilon)$, i.e.,

$$\sum_{|k| \leqq n}{}' c_k e^{i\mu_k x} - \sum_{|k| \leqq n} \omega_k e^{ikx}$$

converges uniformly to zero. Since (108) converges to $f(x)$ in $(-\pi + \epsilon, \pi - \epsilon)$, the theory of trigonometric series ([6], 286(i)) tells us that (108) is uniformly equiconvergent with the Fourier series S of a function of period 2π which equals $f(x)$ in $(-\pi + 2\epsilon, \pi - 2\epsilon)$. Let

$$(109) \qquad f(x) \sim \sum g_k e^{i\mu_k x}, \qquad (-\pi, \pi).$$

By Lemma 1, the series (109) is uniformly equiconvergent with the Fourier series of $f(x)$ in $(-\pi, \pi)$. Hence the series (109) is uniformly equiconvergent with S in $(-\pi + 3\epsilon, \pi - 3\epsilon)$, and hence with (108) and with (107). Since ϵ is arbitrary, it follows that

$$\sum_{|k| \leqq n}{}' c_k e^{i\mu_k x} - \sum_{|k| \leqq n} g_k e^{i\mu_k x}$$

converges to zero for $|x| < \pi$.

Using the notation and results of Section 11, $\sum{}' g_0\gamma_k e^{i\mu_k x}/(1-\gamma_0)$ converges uniformly to g_0 in $(-\pi, \pi)$. Hence

$$\sum{}' [c_k - g_k - g_0\gamma_k/(1-\gamma_0)] e^{i\mu_k x}$$

converges to zero in $(-\pi, \pi)$. By Theorem 9, the desired result follows.

REFERENCES.

[1] U. Dini, *Serie di Fourier e altre rappresentazione analitiche delle funzioni di una variabile reale*, Pisa, 1880, pp. 174-211.

[2] ———, "Sugli sviluppi in serie $\frac{1}{2}a_0 + \sum(a_n \cos \lambda_n x + b_n \operatorname{sen} \lambda_n x)$ dove le λ_n sono radici della equazione trascendente $F(z) \cos \pi z + F_1(z) \operatorname{sen} \pi z = 0$," *Annali di Matematica*, ser. 3, vol. 26 (1917), pp. 261-284.

[3] S. Verblunsky, "On the roots of a transcendental equation which occurs in the theory of trigonometric series," *Mathematische Zeitschrift*, vol. 61 (1954), pp. 324-335.

[4] ———, "Some theorems on F. A. series," *Rendiconti del Circolo Matematico di Palermo*, ser. 2, vol. 3 (1954), pp. 89-105.

[5] F. Wolf, "Contributions to a theory of summability of trigonometric integrals," *University of California Publications in Mathematics*, New Series, vol. 1 (1947), pp. 159-228.

[6] A. Zygmund, *Trigonometrical Series*, Warsaw, 1935.

[7] ———, "On trigonometric integrals," *Annals of Mathematics*, vol. 48 (1947), pp. 393-440.

ISOMORPHISM BETWEEN H^p AND L^p.*

By R. P. Boas, Jr.

Let p be finite and greater than 1. By L^p we understand the Banach space of complex-valued measurable functions f (those differing only on a set of measure zero being identified) of period 2π, such that $\int_0^{2\pi} |f(x)|^p \, dx$ is finite, with norm

$$\| f \| = \{ \int_0^{2\pi} | f(x) |^p \, dx \}^{1/p}.$$

An element f of L^p has a formal Fourier series,

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}, \qquad c_n = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} \, dx. \tag{1}$$

M. Riesz's theorem on conjugate functions may be stated in the following form (cf., in particular, [1], p. 791):

If $f \varepsilon L^p$ *has the Fourier series* (1), *then* (1) *with* c_n *replaced by* 0 *for* $n < 0$ *is the Fourier series of a function* f_+ *of* L^p *and* $\| f_+ \| \leqq A_p \| f \|$, *with* A_p *independent of* f.

By H^p we understand the Banach space, with the L^p norm, whose elements are the functions f_+ obtained from L^p-functions according to this theorem. It is known ([3]; [2], p. 293) that H^p can alternatively be defined as the Banach space (with the natural norm) of boundary values of functions F which are analytic in $| z | < 1$ and satisfy

$$\int_0^{2\pi} | F(re^{i\theta}) |^p \, d\theta \leqq M, \qquad 0 < r < 1,$$

with M independent of r.

Our object is to give a simple proof of the fact, which may be useful in investigations involving the spaces H^p, that *the topological vector-spaces* H^p *and* L^p *are isomorphic.* The question of whether there is an isomorphism

---

* Received March 7, 1955.

between them which preserves the metric is left open, as well as the question of the relationships of H^1 and H^∞ to L^1 and L^∞.

Let f be a given element of L^p with Fourier series (1). Let f_+ be defined as in Riesz's theorem and let $f_- = f - f_+$; both these functions have period 2π. Map f into the function g defined by

$$g(x) = f_+(2x) + e^{-ix} f_-(-2x);$$

then the Fourier series of $g(x)$ is $c_0 + \sum_{n=1}^{\infty} (c_n e^{2inx} + c_{-n} e^{(2n-1)ix})$, so that $g \in H^p$ and $\| g \| \leqq 2A_p \| f \|$. Thus the mapping is bounded; and it is clearly distributive. It is also one-to-one and onto. For, let g be a given element of H^p with Fourier series $\sum_{n=0}^{\infty} b_n e^{inx}$. The series

$$(2) \qquad \sum_{n=0}^{\infty} b_{2n} e^{2inx}, \qquad \sum_{n=1}^{\infty} b_{2n-1} e^{(2n-1)ix}$$

are Fourier series of functions in H^p; indeed, the first is the Fourier series of $g_1(x) = \frac{1}{2}\{g(x) + g(x+\pi)\}$, and the second is then the Fourier series of $g_2(x) e^{-ix}$, where $g_2(x) \sim \sum_{n=1}^{\infty} b_{2n-1} e^{2inx}$. We can map g into the element f of L^p given by $f(x) = g_1(x/2) + g_2(-x/2)$. Applied to the element g obtained from f by the original mapping, this leads back to f; so the inverse mapping from H^p to L^p is one-to-one. Finally, this inverse mapping is bounded, either by another application of Riesz's theorem or because it is the inverse of a one-to-one bounded mapping.

NORTHWESTERN UNIVERSITY.

REFERENCES.

[1] S. Bochner, "Additive set functions on groups," *Annals of Mathematics*, vol. 40 (1939), pp. 769-799.

[2] W. W. Rogosinski and H. S. Shapiro, "On certain extremum problems for analytic functions," *Acta Mathematica*, vol. 90 (1953), pp. 287-318.

[3] A. E. Taylor, "On certain Banach spaces whose elements are analytic functions," *Actas Acad. Ci. Lima*, vol. 12 (1949), pp. 31-43.

ON FINITE GROUPS WITH CYCLIC SYLOW SUBGROUPS FOR ALL ODD PRIMES.*

By MICHIO SUZUKI.

The purpose of this paper is to determine the structure of some finite groups in which all Sylow subgroups of odd order are cyclic. This assumption on Sylow subgroups simplifies the structure of groups considerably, but the structure of 2-Sylow subgroups might be too complicated to make any definite statement on the structure of the groups. In this paper, therefore, we shall make another assumption on 2-Sylow subgroups, and our main result may be stated as follows.

Let G be a non-solvable group of finite order. We assume that all Sylow subgroups of odd order are cyclic, and moreover that a 2-Sylow subgroup is either (a) a dihedral group, or (b) a generalized quaternion group. Then G contains a normal subgroup G_1 such that $[G:G_1] \leqq 2$ and $G_1 = Z \times L$, where Z is a solvable group whose Sylow subgroups are all cyclic, and L is isomorphic with the linear fractional group $LF(2,p)$ over the prime field of characteristic p in the case (a), and with the special linear group $SL(2,p)$ in the case (b).

It has been proved by Frobenius and Burnside (see Burnside [7]) that if all Sylow subgroups of a finite group G are cyclic, then G is solvable. In fact, such a group must be meta-cyclic. Zassenhaus [13] has studied the structure of such groups completely (see Zassenhaus [13], or [14], § 4). In the same paper, Zassenhaus studied the *solvable* groups in which all Sylow subgroups of odd order are cyclic and a 2-Sylow subgroup contains a cyclic subgroup of index 2, and he classified all possible types of such groups. His results are used frequently in the course of the proof of our main result. We refer for these results to the paper [13] of Zassenhaus.

In the first section we shall give a short summary of known properties of linear groups, which are necessary for our purpose, referring mostly to the work by Schur [10]. Some general lemmas on group characters are proved in the next section 2. Our Lemmas 4 and 5 may be regarded as slight generalizations of a result by Brauer and they are main tools for the proof. Section 3 contains several characterizations of linear groups by their group

---

* Received November 11, 1954.

theoretical properties, obtained by Brauer, Suzuki, and Wall. Sections 4 and 5 are devoted to the proof of our main theorem in case of dihedral 2-Sylow subgroups. The case when 2-Sylow subgroups are generalized quaternion groups is treated in Section 6, where we consider a slightly more general situation (see Theorem D). In the proof of Theorem D we shall use very delicate properties of group characters, namely the block theory of Brauer. The final section contains remarks on the applications and the converse of our main theorem.

We use the standard notations. The *linear fractional group* $LF(2,p)$ is the totality of linear fractional transformations $x' = (\alpha x + \beta)/(\gamma x + \delta)$ with elements α, β, γ, δ in the prime field F of characteristic p, and with determinant $\alpha\delta - \beta\gamma = 1$. The *special linear group* $SL(2,p)$ is the totality of non-singular 2×2 matrices over F with determinant 1. A *generalized quaternion group* is a non-cyclic 2-group which contains only one subgroup of order 2. It is generated by two elements σ and ρ with the relations

$$\sigma^{2^{n+1}} = 1, \quad \rho^2 = \sigma^{2^n} \quad \text{and} \quad \rho\sigma\rho^{-1} = \sigma^{-1}.$$

The central factor group of a generalized quaternion group is a *dihedral group.* A finite group is called a *Z-group* if all Sylow subgroups are cyclic. For the properties of Z-groups see Zassenhaus [13]. Group elements are denoted by Greek letters σ, ρ and τ; τ stands usually, not exclusively, for elements of order 2. An element σ is called *p-regular* if the order of σ is prime to p. Otherwise σ is *p-singular.*

1. We shall make preliminary remarks on the special linear group $SL(2,p)$ and the linear fractional group $LF(2,p)$ over the prime field of characteristic p.

The group $SL(2,p)$ contains only one element of order 2, which generates the center, and the central factor group is isomorphic with $LF(2,p)$. It has been proved by Schur, [10], p. 120, that if a finite group G contains a normal subgroup Z of order 2, and if the factor group G/Z is isomorphic with $LF(2,p)$, then G is either isomorphic with $SL(2,p)$, or is a direct product of Z and $LF(2,p)$. In particular we have

LEMMA 1. *If a finite group G contains only one subgroup Z of order 2, and if the factor group G/Z is isomorphic with $LF(2,p)$, then G is isomorphic with $SL(2,p)$.*

We shall consider the group of all automorphisms of $LF(2,p)$. The following lemma is proved.

LEMMA 2. *The index of the group of inner automorphisms in the group of all automorphisms of $LF(2,p)$ is two.*

Proof. Let L be the totality of linear fractional transformations over the $GF(p)$. The transformations with determinant 1 form a normal subgroup of index 2, isomorphic with $LF(2,p)$. Now let A be the group of all automorphisms of $LF(2,p)$. Then A contains a subgroup isomorphic with L. $LF(2,p)$ may be regarded as a subgroup I of A, consisting of all inner automorphisms. I contains $p+1$ subgroups of order p, and A may be regarded as a substitution group on these subgroups. Since $A \supseteq L$, A is at least triply transitive. Suppose that an element σ leaves invariant three subgroups P_1, P_2 and P_3. Consider the normalizer N_i of P_i in I. Since σ leaves invariant P_i, every N_i is also fixed by σ, and so are $N_1 \cap N_2$ and $N_1 \cap N_3$. Now N_1 may be generated by two elements τ and ρ such that

$$\{\tau\} = P_1, \qquad \{\rho\} = N_1 \cap N_2, \qquad \{\tau\rho\} = N_1 \cap N_3$$

and

$$\rho\tau\rho^{-1} = \tau^r.$$

Since σ maps $\{\tau\}$, $\{\rho\}$ and $\{\tau\rho\}$ into themselves, we have

$$\sigma\tau\sigma^{-1} = \tau^m, \quad \sigma\rho\sigma^{-1} = \rho^n \quad \text{and} \quad \sigma(\tau\rho)\sigma^{-1} = (\tau\sigma)^k,$$

where m, n and k are positive integers. These equations imply that $\rho^n\tau^m\rho^{-n} = \tau^{mr}$ or $mr \equiv mr^n \pmod{p}$. Since $r^{(p-1)/2} \equiv 1 \pmod{p}$, we have $n \equiv 1 \pmod{(p-1)/2}$. Moreover since $k \equiv n \pmod{(p-1)/2}$, we may assume that $n = k = 1$, and hence $m = 1$. Thus σ induces the identical automorphism on N_1, and similarly on N_2. Hence σ transforms all elements of I into themselves, and so is in the centralizer of I in A. On the other hand, the centralizer of I in A contains only the identity, since the center of I is trivial. This proves $\sigma = 1$. Hence A is representable as a triply transitive permutation group on $p+1$ letters such that no element except the identity leaves invariant three letters. By a theorem of Zassenhaus [12], A is isomorphic with a linear fractional group, and so A coincides with L. This proves Lemma 2.

We shall now consider the irreducible characters of $SL(2,p)$. The irreducible characters of $SL(2,p)$ have been determined by Schur [10] (see in particular the table on p. 128) and independently by Jordan [9].

Let G be the special linear group $SL(2,p)$ over the prime field of characteristic p. We are interested in the structure of 2-blocks of characters of G. For the theory of blocks see [1], [2] and [3]. Let ϵ be 1 or -1, so that $p \equiv \epsilon \pmod 4$, and let 2^α be the highest power of 2 which divides $p - \epsilon$:

$$p - \epsilon = 2^\alpha h, \qquad (h, 2) = 1 \qquad\qquad (\alpha \geqq 2).$$

If we set $(p + \epsilon)/2 = 2a + 1$, G has exactly a 2-blocks of defect 1. This may be proved either directly from the table of characters, or by using a result of Brauer [3], § 8, since the factor group G/Z modulo the center Z of G is isomorphic with $LF(2, p)$, and $LF(2, p)$ has exactly a 2-blocks of defect 0. From the general theory of Brauer [1], each 2-block of defect 1 consists of two characters, and they take the same value on every 2-regular element. The degree of characters in a block of defect 1 is $p - \epsilon$. Some characters of degree $p + \epsilon$ may belong to blocks of defect α. In fact, if we put $h - 1 = 2b$, G has exactly b 2-blocks of defect α. From the table of characters we see that each block of defect α contains 2^α characters which take the same values on every 2-regular element.

The first 2-block B_1 of G is the only block of defect $\alpha + 1$. It consists of the principal character 1, one character ϕ_1 of degree p, two p-conjugate characters ϕ_2 and ϕ_3 of degree $(p + \epsilon)/2$, which belong to the representations of the central factor group, two conjugate characters ϕ_4 and ϕ_5 of degree $(p - \epsilon)/2$, and $2^{\alpha-1} - 1$ characters ϕ_i $(i = 6, 7, \cdots, 2^{\alpha-1} + 4)$ of degree $p + \epsilon$.

There are several linear combinations of these characters which vanish on all 2-regular elements. The following table may be obtained from the table of characters (using suitable notations).

Relations \ on elements of order	prime to 2	$p + \epsilon$	2^α	2
$1 + \epsilon\phi_1 - \epsilon\phi_2 - \epsilon\phi_3$	0	0	4	0
$1 - \epsilon\phi_2 + \epsilon\phi_4$	0	2	2	$1 - \epsilon p$
$1 - \epsilon\phi_3 + \epsilon\phi_5$	0	2	2	$1 - \epsilon p$
$\phi_i - \phi_2 - \phi_3$	0	0	$2\epsilon + \epsilon\theta_i$	0 $-2(p + \epsilon)$

TABLE I

In the last line of Table I, θ_i $(i = 6, 7, \cdots, 2^{\alpha-1} + 4)$ are the values of non-linear characters of a generalized quaternion group of order $2^{\alpha+1}$, taken in a suitable arrangement; and, in the lower right hand corner, the value 0 is for those ϕ_i which belong to representations of the central factor group and the value $-2(p + \epsilon)$ is for those which do not.

2. In this section we shall prove some lemmas on group characters. Some of these lemmas have been obtained by Brauer and Fowler [5].

Let $C_1, C_2, \cdots, C_k$ be conjugate classes of a finite group G of order g, and $s_1, \cdots, s_k$ the totality of distinct irreducible characters. The product C_iC_j may be written as a sum of conjugate classes: $C_iC_j = \sum_{l=1}^{k} a_{ijl}C_l$, where a_{ijl} are non-negative integers. The following formula is a consequence of orthogonality relations (see [5]):

$$(*) \qquad a_{ijl} = g_ig_jg^{-1} \sum_{\mu} z_\mu^{-1} s_\mu(C_i) s_\mu(C_j) \bar{s}_\mu(C_l).$$

In this formula, g_i is the number of elements in C_i, z_μ is the degree of s_μ and $\bar{s}_\mu$ indicates the conjugate complex of s_μ. In the application to our case, classes C_i and C_j contain elements of order 2. We have

LEMMA 3. *Let G be a finite group of even order. Assume that G has only one class T which contains an element of order 2, and that $T^2 = \sum_{l=1}^{k} a_lC_l$. Then we have*

$$a_l = gn^{-2} \sum_{\mu} z_\mu^{-1} s_\mu(T)^2 s_\mu(C_l)$$

where n is the order of the normalizer of an element of order 2. Moreover if the class C_l is not self-reciprocal, then $a_l = 0$.

Proof. Let σ and τ be two elements of order 2. Then the inverse of $\sigma\tau$ is $\tau\sigma$, which is obviously conjugate to $\sigma\tau$. This proves the last assertion. The first assertion is a consequence of (*).

By a (generalized) character, we shall mean a linear combination of irreducible characters with rational integral coefficients. Let ϕ be a character of G. Then $\phi = \sum_\mu a_\mu s_\mu$, and all a_μ are rational integers, uniquely determined by ϕ; namely $a_\mu = \int_G \phi\bar{s}_\mu$. (The symbol $\int_G$ means the average on G, i.e. the summation over all elements of G divided by the order of G.) The sum of the square of these coefficients is denoted by $w(\phi)$, and may be calculated as $w(\phi) = \sum_\mu a_\mu^2 = \int_G |\phi|^2$, where $|\phi|$ is the absolute value of ϕ.

We consider a subgroup H of G and a character ψ of H. The induced character ψ^* of G may be defined by

$$\psi^*(\sigma) = h^{-1} \sum_{\rho \varepsilon G} \psi(\rho\sigma\rho^{-1}),$$

where h is the order of H, and we have put $\psi(\sigma) = 0$ if $\sigma \notin H$. Clearly, ψ^*

is a character of G (in our sense). If $\psi = \sum a_i\psi_i$, then we have $\psi^* = \sum a_i\psi_i^*$.

A conjugate class C of a subgroup H is called *special* (with respect to G), if C contains an element σ such that the centralizer of σ in G is contained in H. We shall prove the following lemma.

LEMMA 4. *Let $C_1, C_2, \cdots, C_r$ be special classes of H, and $\sigma_i \varepsilon C_i$ $(i = 1, 2, \cdots, r)$. Assume that σ_i is not conjugate to σ_j even in G if $i \neq j$. If ψ is a character of H which vanishes outside of these C_i, then $\psi^*(\sigma_i) = \psi(\sigma_i)$ $(i = 1, 2, \cdots, r)$, and $\psi^*(\sigma) = 0$ if σ is not conjugate to any σ_i. Moreover we have $w(\psi^*) = w(\psi)$.*

Proof. By the definition of induced character,

$$\psi^*(\rho) = h^{-1} \textstyle\sum_{\tau \varepsilon G} \psi(\tau\rho\tau^{-1}).$$

Hence if $\psi^*(\rho) \neq 0$, at least one term of the right hand side does not vanish. But ψ is zero outside of C_i, and so $\psi^*(\rho) \neq 0$ implies that ρ is conjugate to some σ_k. This proves the second assertion.

If $\rho = \sigma_i$, then $\psi(\tau\rho\tau^{-1}) \neq 0$ only when $\tau\rho\tau^{-1} \sim \sigma_k$ for some k. By our assumption σ_i is conjugate to σ_k only when $k = i$. Hence $\psi(\tau\rho\tau^{-1}) \neq 0$ implies that $\tau\rho\tau^{-1}$ is conjugate to σ_i in H, i.e. $\tau\rho\tau^{-1} = \nu\sigma_i\nu^{-1}$ $(\nu \varepsilon H)$. Then $\nu^{-1}\tau$ is in the normalizer of σ_i. Since C_i is special, the normalizer of σ_i is contained in H. Hence we conclude that $\tau \varepsilon H$. We see now that

$$\psi^*(\sigma_i) = h^{-1} \textstyle\sum_{\tau \varepsilon H} \psi(\tau\sigma_i\tau) = \psi(\sigma_i).$$

Let N_i be the normalizer of σ_i in G. By our assumption, we may take suitable σ_i so that N_i is a subgroup of H. The number of elements which are conjugate to σ_i in G is denoted by g_i, and similarly h_i is the number of elements in C_i. Then $g_i = [G : N_i]$ and $h_i = [H : N_i]$. Hence we have $g_i = [G : H]h_i$ $(i = 1, 2, \cdots, r)$. We shall now compute $w(\psi^*)$. We have

$$\begin{aligned} w(\psi^*) &= \int_G |\psi^*|^2 = g^{-1} \sum_i g_i |\psi^*(\sigma_i)|^2 \\ &= g^{-1} \sum_i [G : H]h_i |\psi(\sigma_i)|^2 = h^{-1} \sum_i h_i |\psi(\sigma_i)|^2 = \int_H |\psi|^2 = w(\psi), \end{aligned}$$

since $\psi = 0$ outside of C_i, which proves the statement.

As an application we shall prove the following lemma, due essentially to Brauer (see [4]).

LEMMA 5. *Let $C_1, \cdots, C_r$ be special classes of H satisfying the same assumption as in Lemma 4. If H has s irreducible characters $\phi_1, \cdots, \phi_s$ such that $\phi_i = \phi_j$ outside of C_k, and $\phi_i(\sigma) = 0$ for any element σ of H which* *is conjugate to σ_k in G but not in H, and if $s \geqq 2$, then G has exactly s irreducible characters $s_1, \cdots, s_s$ such that*

$$\phi_i^* = \epsilon s_i + a \sum_{k=1}^{s} s_k + \Delta, \qquad (i = 1, 2, \cdots, s),$$

where $\epsilon = \pm 1$, $a \geqq 0$, $a + \epsilon \geqq 0$ and Δ is a character of G which is independent of i and does not contain any s_j. All characters of G other than these s_i contain each ϕ_k with the same multiplicity (as characters of H).

Definition. These characters $s_1, \cdots, s_s$ are called the *exceptional* characters associated with $\phi_1, \cdots, \phi_s$.

Proof. Since $s \geqq 2$, we may consider $s - 1$ characters $\psi_i = \phi_i - \phi_s$ $(i = 1, 2, \cdots, s - 1)$ of H which satisfy the assumption of Lemma 4. Hence $w(\psi_i^*) = w(\psi_i) = 2$, and so each induced character ψ_i^* contains exactly 2 different irreducible characters of G with multiplicity $+1$ or -1. Since

$$w(\psi_i^* - \psi_j^*) = w(\psi_i - \psi_j) = 2 \qquad (i \neq j),$$

and since $\psi_i^* = 0$ for the identity of G, we may write $\psi_i^* = \epsilon(s_i - s_s)$, $(i = 1, 2, \cdots, s - 1)$, with $\epsilon = +1$ or -1. Hence we have $\phi_i^* - \epsilon s_i = \phi_s^* - \epsilon s_s$. This implies that ϕ_i^* may be decomposed in the following manner:

$$\phi_i^* = \epsilon s_i + \sum_{j=1}^{s} a_j s_j + \Delta,$$

where Δ is either identically 0 or a sum of irreducible characters of G other than $s_1, \cdots, s_s$. Consider now the difference $w(\phi_i^*) - w(\phi_s^*)$. This is equal to

$$g^{-1} \sum_{\sigma \varepsilon G} (\phi_i^*(\sigma)\phi_i^*(\sigma^{-1}) - \phi_s^*(\sigma)\phi_s^*(\sigma^{-1})).$$

If σ is not conjugate to any σ_k $(k = 1, 2, \cdots, r)$, then $\phi_i^*(\sigma) = \phi_s^*(s)$. If σ is conjugate to σ_k, then from our assumption on ϕ_i we see that $\phi_i^*(\sigma_k) = \phi_i(\sigma_k)$. Hence as in the proof of Lemma 4, we conclude that

$$w(\phi_i^*) - w(\phi_s^*) = w(\phi_i) - w(\phi_s) = 0.$$

Hence we have $(a_i + \epsilon)^2 + a_s^2 = a_i^2 + (a_s + \epsilon)^2$, or $a_i = a_s$. Hence we obtain $\phi_i^* = \epsilon s_i + a \sum_{k=1}^{s} s_k + \Delta$, as claimed.

3. In the proof of our main theorems, we shall need some group theoretical characterization of $LF(2, p)$. We shall state them in the following.

LEMMA 6. *Let G be a finite simple group of even order, satisfying the following property: If two cyclic subgroups U and V have a non-trivial intersection, i.e., $U \cap V \neq e$, and if at least one of these subgroups has an even order, then $U \cup V$ is cyclic. Then G is isomorphic with $LF(2,q)$ for some prime power q.*

This result is due to Brauer. See [6]. From this lemma, we conclude the following

LEMMA 7. *Let G be a simple group of finite even order. We assume that all Sylow subgroups of odd order are cyclic, and that the normalizer of an element of order 2 is of dihedral type. Then G is isomorphic with $LF(2,p)$ for some prime p.*

Proof. Let τ be an element of order 2, of which the normalizer N is of dihedral type. Put $T = \{\tau\}$. Consider a 2-Sylow subgroup S of N. S is contained in a 2-Sylow subgroup S^* of G. If $S \neq S^*$, T is not a characteristic subgroup of S, and this implies that S is of order 4. Since S is the normalizer of T in S^*, S^* is generated by two elements σ and τ such that $\sigma^{2^n} = 1$, $\tau\sigma\tau^{-1} = \sigma^r$, where $r = -1$ or $-1 + 2^{n-1}$. (See Lemma 4 of [11]). Hence τ is not conjugate to the central element of S^* in G. By a theorem of Grün [8], G is not perfect. This contradicts our assumption. Hence $S = S^*$ is a 2-Sylow subgroup of G. We see now that the totality of elements of order 2 forms a single conjugate class of G. This follows also from the theorem of Grün (loc. cit.).

Now N contains a normal subgroup Z such that $ZS = N$ and $Z \cap S = e$. If $Z = e$, then G satisfies the assumption of Lemma 6. Hence our assertion follows immediately.

Assume that $Z \neq e$. Consider any subgroup U of Z. The centralizer V of U in G has a cyclic 2-Sylow subgroup. Hence V is a Z-group. We shall prove that $V \subseteq N$ using induction on the index $[Z:U]$. First of all assume that $U = Z$. Suppose $V \not\subseteq N$. Denoting by M the maximal cyclic subgroup of N with index 2, V contains a normal subgroup W such that $WM = V$ and $W \cap M = Z$. In the factor group V/Z, the normalizer of TZ/Z is M/Z. Hence the element τ of T induces an automorphism f in W/Z which maps every element onto its inverse. Now W is a normal subgroup of the normalizer of Z in G, which contains an element σ of order 2, $\sigma \notin T$. This element σ induces an automorphism of each Sylow subgroup of W/Z, which is either the identical mapping or the same one as f. This implies that every prime divisor of $[W:Z]$ also divides the order of Z, since any element of order 2 is conjugate to τ. On the other hand the order of the commutator subgroup of a Z-group is prime to the index of commutator group, which is a multiple of the order of its center. Hence the commutator subgroup of V must be e, and so V is cyclic. Since $V \supseteqq T$, and we get a contradiction. Hence the centralizer of Z is contained in N.

Now we consider any subgroup U of Z. Consider all subgroups U_i of Z such that $U_i \supset U$ and U_i/U are Sylow subgroups of Z/U. By the inductive hypothesis, each U_i commutes only with elements of N, since the centralizer of U_i is M, and so the normalizer of U_i is contained in the normalizer N of T. Suppose that the centralizer V of U is not contained in N. In V/U, each U_i/U is a Sylow subgroup and is contained in the center of its normalizer. Hence by a theorem of Burnside (see e.g. [14], p. 133), V contains a normal subgroup W such that $WM = V$ and $W \cap M = U$. We may apply the same argument as before to get a similar contradiction.

We have thus proved that the centralizer of any subgroup of Z coincides with M. Hence our group G satisfies the property of Lemma 6, and is isomorphic with $LF(2, p)$ for some prime number p. This proves Lemma 7.

Finally we shall state the following result (see [6]).

LEMMA 8. *Let G be a non-solvable group of finite even order. If G contains an abelian subgroup A of even order such that $A \neq e$ and the centralizer of any element $\neq 1$ of A is contained in A, then G is isomorphic with $LF(2, 2^{\mu})$.*

4. This section is devoted to the proof of the following theorem:

THEOREM A. *Let G be a non-solvable group of even order. We assume that all l-Sylow subgroups for odd l are cyclic, and that 2-Sylow subgroups are of order 4. Then G is a direct product of $LF(2, p)$ and a Z-group.*

Proof. We shall use an induction on the order g of G.

First of all, suppose that the commutator subgroup G' of G is a proper subgroup of G. Then G' satisfies all assumptions of this Theorem A, and has a smaller order than G. We may, therefore, apply our inductive hypothesis to G'. G' is a direct product of a Z-group U and a subgroup L, isomorphic with $LF(2, p)$ for some p. Since U is solvable, L is a characteristic subgroup of G', and hence L is a normal subgroup of G. Consider now the centralizer V of L in G. V is a normal subgroup of G, and the factor group G/V is isomorphic with a subgroup of the group L_0 of all automorphisms of L. By Lemma 2, the index $[L_0 : L]$ is 2. On the other hand, by assumption the order of G is not divisible by 8, and hence $[G: V \cup L]$ is odd. This implies that $G = V \cup L$, or $G = V \times L$.

We may therefore assume that $G = G'$, i.e. that G is perfect. Suppose that G contains a proper normal subgroup H. If $[G:H]$ were not divisible by 4, then the factor group G/H would be a Z-group, and so solvable. This is not the case. If $[G:H] \equiv 0 \ (4)$, then by assumption H would be of odd order, and so a Z-group. Hence H would contain a characteristic subgroup of odd prime order q. This is however impossible, because no perfect group with cyclic q-Sylow subgroups can contain a normal subgroup of order q by Theorem 15 of Zassenhaus [13]. Hence G must be a simple group.

Let S_2 be a 2-Sylow subgroup of G, and let Z_2 be the centralizer and N_2 the normalizer of S_2 in G. Since G is simple, $N_2 \neq Z_2$. Hence S_2 is not cyclic, and $[N_2 : Z_2] = 3$. Z_2 is a direct product of S_2 and a Z-group U: $Z_2 = U \times S_2$.

Assume that $U = e$. Take a subgroup T of order 2 in S_2. Consider the normalizer N of T in G. The factor group N/T is a Z-group, and hence N contains a normal subgroup Z such that $N = ZS_2$ and $Z \cap S_2 = e$. An element of S_2, which is not in T, induces an automorphism of Z, and this automorphism leaves only the identity invariant. Hence by a lemma of Zassenhaus ([12], Lemma 1) Z is a cyclic group and N is a dihedral group. Theorem A now follows from Lemma 7.

Suppose finally that $U \neq e$. This case will be proved impossible at the end of this section. Again denote by T a subgroup of order 2 in S_2. We consider the structure of the normalizer N of T in G.

Consider the normalizer $N(U)$ of U in G. Suppose $N(U) \neq N_2$. If the factor group $N(U)/U$ is solvable, then by Theorem 6 of Zassenhaus [13], $N(U)$ contains a normal subgroup V such that $VN_2 = N(U)$ and $V \cap N_2 = U$. Since V is of odd order, V/U is a Z-group, and hence V contains a subgroup P such that P is a normal subgroup of $N(U)$ and $[P:U] = l$ is a prime number. Then the centralizer C/U of P/U contains Z_2/U, since $N(U)/C$ is cyclic. This implies that P is a part of the centrailzer of S_2. This gives a contradiction. Hence $N(U)/U$ is not solvable. We may apply the inductive hypothesis to $N(U)$ since $G \neq N(U)$. It follows that $N(U)$ is a direct product of U and a subgroup L, which is isomorphic with some $LF(2,p)$. Let q be the smallest prime divisor of the order of U, Q the q-Sylow subgroup of U. The normalizer $N(Q)$ of Q in G contains L, and so it is not solvable. By the inductive hypothesis $N(Q) = L' \times U'$ where $L' \supseteq L$ and $U' \subseteq U$. Since q is the smallest prime divisor of the order of U, $N(Q)$ coincides with the centralizer of Q. This implies that a q-Sylow subgroup of G is in the center of its normalizer. By a theorem of Burnside (loc. cit.), G is not simple. This contradiction comes from the assumption that $N(U) \neq N_2$, and hence $N(U) = N_2$.

Consider next a maximal subgroup M of G, which contains N_2. If M is not solvable, $N(U)$ would be larger than N_2 by the inductive hypothesis. M is, therefore, solvable. Hence by Theorem 6 of Zassenhaus [13], M contains a normal subgroup K such that $M = KN_2$ and $K \cap N_2 = U$. Since K is a Z-group, the commutator subgroup K' of K is cyclic. Take a Sylow subgroup P of K'. P is normal in M, so that the centralizer $Z(P)$ is also normal, and $M/Z(P)$ is cyclic. Hence $Z(P)K$ contains S_2, and this shows that $Z(P) \supseteq S_2$, or $P \subseteq Z_2$. Since P was any Sylow subgroup of K', we conclude that $K' \subseteq U$. This implies that U is a normal subgroup of M, or $M \subseteq N(U) = N_2$. In other words, N_2 is a maximal subgroup of G.

We shall show that U is cyclic. N_2 contains a subgroup W such that $WS_2 = N_2$ and $W \cap Z_2 = U$. Suppose that the commutator subgroup W' of W is a proper part of U. We consider any prime divisor of $[U : W']$, say q, and the q-Sylow subgroup Q of U. The normalizer $N(Q)$ of Q in G is solvable, since otherwise G would be non-simple by Burnside's theorem. Since $N_2 = U(N_2 \cap N(Q))$, $N(Q)$ has no normal subgroup of index 2. Hence again by Theorem 6 of Zassenhaus [13], $N(Q)$ contains a normal subgroup X such that $N(Q)/X$ is isomorphic with the tetrahedral group. Let Y be any Sylow subgroup of X. Then the order of the centralizer of Y in $N(Q)$ is divisible by 4, and hence some conjugate subgroup Y_0 of Y in $N(Q)$ is contained in Z_2, or $X \subseteq U$. Hence $N(Q)$ coincides with the centralizer of Q, which is impossible by virtue of Burnside's theorem. Hence we have $W' = U$, or U is cyclic.

Consider again the normalizer N of the subgroup T of order 2. Suppose that $N \subseteq N_2$. Then Z_2 is abelian, and the centralizer of any element $\neq 1$ of Z_2 coincides with Z_2. By Lemma 8 this is impossible, since we must have $U = e$ in $LF(2, 2^\mu)$.

Hence we consider the case $N \not\subseteq N_2$. We then have $N \cap N_2 = Z_2$, so that we see from Burnside's theorem (loc. cit.) that N contains a normal subgroup H such that $HS_2 = N$ and $H \cap S_2 = e$. If H' is the commutator subgroup of H, the subgroup $H' \cap U$ is a normal subgroup of both N and N_2. Since N_2 is maximal, we must have $H' \cap U = e$. H contains a cyclic subgroup X such that $X \supseteq U$, $X \cap H' = e$ and $XH' = H$. Since the normalizer of U is N_2, $X \subseteq N_2$ or $X = U$. In particular, we have $UH' = H$. The above consideration is applied to any maximal subgroup M of G which contains N. Corresponding to H, M contains a normal subgroup K such that $K \supseteq H$,

$K \cap S_2 = e$ and $KS_2 = M$. We want to show that $M = N$. If $N(H)$ is the normalizer of H in G, $N(H) = H \cup X$, where X is the normalizer of some Sylow subgroup Q of U in $N(H)$. But $N(H) \supseteq N_2$, so that $X = Z_2$, or $N(H) = HZ_2 = N$. Hence if $M \neq N$, the commutator subgroup K' of K is not a part of H. The element of order 2 in T induces an automorphism of $K'/K' \cap H$, which sends every element into its inverse. This implies that $[K' : K' \cap H]$ is prime to $[K' \cap H : e]$ since every Sylow subgroup of K' is cyclic. On the other hand, if q is any prime divisor of $[K' : K' \cap H]$, M contains an element of order $2q$, since the centralizer of a q-Sylow subgroup of K' is of even order. Hence $[K' : K' \cap H]$ divides the order of H, and therefore $[K' : K' \cap H]$ divides $[H : K' \cap H] = [HK' : K']$. This is a contradiction, since the order and index of K' in K must be relatively prime. We must have $M = N$; in other words, N is a maximal subgroup of G.

Let $[U : e] = u$ and $[H' : e] = v$. Then we have $g = [G : e] = 4uvm$, and u, v, m are relatively prime. By assumption $U \neq e$, or $u > 1$.

By a well-known argument of Burnside we see that two elements in H' are conjugate in G if and only if they are conjugate in N. Since N is a maximal subgroup of G and since G is simple, every conjugate class of N containing elements $\neq 1$ of H' is special with respect to G. N has $n = (v - 1)/u$ non-linear irreducible characters of degree $2u$, which vanish on elements outside of TH'. Denote them by $\phi_1, \phi_2, \cdots, \phi_n$. For these characters the assumptions of Lemma 5 are satisfied. Hence we conclude that G has n "exceptional" characters $E_1, \cdots, E_n$ such that $\phi_i^* = \delta E_i + \Delta$, where $\delta = \pm 1$ and Δ is independent of i.

Similarly every conjugate class of N_2 containing elements of Z_2 with order greater than 2 is special. N_2 has $t = (u - 1)/3$ characters $\theta_1, \theta_2, \cdots, \theta_t$ of degree 3, which are the characters of N_2/S_2, and u characters $\xi_1, \cdots, \xi_u$ of degree 3, which belong to representations which have kernels contained in U. All characters θ_i and ξ_j vanish outside of Z_2. Hence by Lemma 5, G has exceptional characters Θ_i and Ξ_j associated with θ_i and ξ_j respectively. (Note that $t \geqq 2$, since u is odd.) As a character of N_2 Ξ_j contains each θ_i with the same multiplicity, but Θ_k does not. Hence $\Xi_j \neq \Theta_k$.

Let $\bar{\theta}_i$ be the linear character of N whose restriction θ_i on Z_2 induces θ_i in N_2. Then we have $\theta_i^* = \bar{\theta}_i^* = \theta_i^* + \sum' \phi_k^*$, where $\sum'$ denotes the summation over all subscripts k for which ϕ_k belongs to a representation of N/T. Hence for $i \neq j$ we have

$$\bar{\theta}_i^* - \bar{\theta}_j^* = \theta_i^* - \theta_j^* = \epsilon_1 (\Theta_i - \Theta_j),.$$

or $\bar{\theta}_i^* = \epsilon_1 \Theta_i + \Delta_1$, $(i = 1, 2, \cdots, t)$, where $\epsilon_1 = \pm 1$ and Δ_1 is independent of i. Similarly for the principal characters $\bar{\theta}$ of N and $\tilde{\theta}$ of Z_2, we have $\bar{\theta}^* = \tilde{\theta}^* + \sum' \phi_k^*$. We apply Lemma 4 to the character $\psi = 1 + \omega + \omega' - \theta_i$ of N_2, where $1, \omega, \omega'$ are linear characters of N_2/Z_2. Then $w(\psi^*) = w(\psi) = 4$. Since ψ^* contains the principal character of G, ψ^* must be a sum of four irreducible characters. Hence we have

$$\bar{\theta}^* - \bar{\theta}_i^* = \tilde{\theta}^* - \tilde{\theta}_i^* = 1 + \delta_1 X_1 + \delta_2 X_2 + \delta_3 X_3 \qquad (\delta_i = \pm 1).$$

One of these characters X_i must be Θ_i, and we have say $X_3 = \Theta_i$ and $\delta_3 = -\epsilon$. Hence $\bar{\theta}^* = 1 + \delta_1 X_1 + \delta_2 X_2 + \Delta_1$.

We shall consider the linear character $\bar{\xi}_j$ of N such that $\bar{\xi}_j(\tau) = -1$ for the generator τ of T, and the restriction $\tilde{\xi}_j$ of $\bar{\xi}_j$ on Z_2 induces ξ_j of N_2. By the same argument as above, we conclude that

$$\xi_j^* = \tilde{\xi}_j^* = \bar{\xi}_j^* + \sum'' \phi_k^*, \text{ and } \bar{\xi}_j^* = \epsilon_2 \Xi_j + \Delta_2. \qquad (\epsilon_2 = \pm 1),$$

where $\sum''$ extends over all characters ϕ_k for which $\phi_k(\tau) = -2u$, and Δ_2 is independent of j. Similarly if $\bar{\xi}_j'$ is the linear character of N such that $\bar{\xi}_j'(\tau) = 1$ and its restriction on Z_2 induces ξ_j of N_2, we have

$$\xi_j^* = \bar{\xi}_j'^* + \sum' \phi_k^*, \text{ and } \bar{\xi}_j'^* = \epsilon_3 \Xi_j + \Delta_3 \qquad (\epsilon_3 = \pm 1).$$

Let η and ϕ be characters of $H'T/T$, and assume that η is the principal character. Then we have

$$\begin{aligned} \eta^* &= \bar{\theta}^* + 3 \sum \bar{\theta}_i^* + \sum \bar{\xi}_j'^* \\ &= 1 + \delta_1 X_1 + \delta_2 X_2 + 3\epsilon_1 \sum \Theta_i + \epsilon_3 \sum \Xi_j + u(\Delta_1 + \Delta_3), \\ \phi^* &= 2 \sum \bar{\xi}_j^* = 2\epsilon_2 \sum \Xi_j + 2u\Delta_2. \end{aligned}$$

The character $\eta - \phi$ of $H'T$ induces a character ψ_0 of N. Since ψ_0 vanishes outside of $H'T$, we may apply Lemma 4. Since $\psi_0^* = \eta^* - \phi^*$, we get $w(\eta^* - \phi^*) = w(\psi_0) = 4u$. Using Lemma 4 we see that

$$\int_G (\bar{\theta}^* - \bar{\theta}_i^*)(\bar{\Xi}_j - \bar{\Xi}_k) = \pm \int_{N_2} \psi(\bar{\xi}_j - \bar{\xi}_k) = 0.$$

Hence X_1 and X_2 are different from Ξ_j, and hence the contributions to $w(\eta^* - \phi^*)$ by the characters $1, X_1, X_2, \Theta_i, \Xi_j$ are at least $1, 1, 1, 9(u-1)/3$, u respectively. Since $1 + 1 + 1 + 3(u-1) + u = 4u = w(\eta^* - \phi^*)$, these values are exact contributions. Hence we conclude that $\epsilon_2 = \epsilon_3$ and $2\Delta_2 = \Delta_1 + \Delta_3$. Comparing the degrees of η^* and ϕ^*, we get

$$2\epsilon_2 \sum_j \mathcal{D}_g \Xi_j = 1 + \delta_1 \mathcal{D}_g X_1 + \delta_2 \mathcal{D}_g X_2 + 3\epsilon_1 \sum_i \mathcal{D}_g \Theta_i + \epsilon_3 \sum_j \mathcal{D}_g \Xi_j.$$

Each Ξ_j has the same degree, and $\mathcal{D}_g\Theta_i$ is also independent of i. Hence using the facts that $\epsilon_2=\epsilon_3$ and $1+\delta_1\mathcal{D}_gX_1+\delta_2\mathcal{D}_gX_2=\epsilon_1\mathcal{D}_g\Theta_i$, we have $\epsilon_2\mathcal{D}_g\Xi_j=\epsilon_1\mathcal{D}_g\Theta_i$, or $\epsilon_1=\epsilon_2$ and $\mathcal{D}_g\Theta_i=\mathcal{D}_g\Xi_j$. If π is any character of Z_2, π^* is either $\bar{\theta}^*$, θ_i^*, or ξ_j^*. Hence the degrees of 1, X_1, X_2, Θ_i, Ξ_j are not divisible by u, but degrees of all other characters of G are multiples of u. Since $\Theta_i=\Theta_j$ on every element of $H'T$, Θ_i is not an exceptional character associated with ϕ_k. Similarly each Ξ_j is not. If X_1 is one of the exceptional characters associated with the ϕ_k of N, then only X_1 and X_2 are exceptional with respect to these ϕ_k. Hence $(v-1)/u=2$, or $v=2u+1$. Since $u\equiv 1 \pmod 3$, v must be divisible by 3. This is impossible, because any prime divisor of v is larger than u. Hence X_1 and X_2 are non-exceptional for ϕ_k.

Since $\xi_j^*=\bar{\xi}_j^*+\sum''\phi_k^*=\bar{\xi}_j'^*+\sum'\phi_k^*$ we have now

$$\epsilon_2\Xi_j+\Delta_2+\delta\sum''E_k+\tfrac{1}{2}n\Delta=\epsilon_3\Xi_j+\Delta_3+\delta\sum'E_k+\tfrac{1}{2}n\Delta,$$

or $\Delta_2+\delta\sum''E_k=\Delta_3+\delta\sum'E_k$. Hence $\Delta_3=\Delta_2-\delta\sum'E_k+\delta\sum''E_k$, and since $\Delta_1+\Delta_3=2\Delta_2$,

$$\Delta_1=\Delta_2+\delta\sum'E_k-\delta\sum''E_k.$$

Now

$$\begin{array}{llr}
\bar{\theta}^*=1+\delta_1X_1+\delta_2X_2 & +\,\delta\sum'E_k-\delta\sum''E_k & +\,\Delta_2,\\
\bar{\theta}_i^*=\epsilon\Theta_i & +\,\delta\sum'E_k-\delta\sum''E_k & +\,\Delta_2,\\
\bar{\xi}_j^*=\epsilon\Xi_j & + & \Delta_2,\\
\bar{\xi}_j'^*=\epsilon\Xi_j & -\,\delta\sum'E_k+\delta\sum''E_k & +\,\Delta_2,\\
\phi_k^*=\delta E_k & & +\,\Delta\ ,
\end{array}$$

where we put $\epsilon_1=\epsilon_2=\epsilon_3=\epsilon$.

Let τ be the generator of T, and ρ the generator of UT. We can compute the values of characters for τ and ρ. This is shown in table II.

Denote by $\langle\tau\rangle$ and $\langle\rho\rangle$ the conjugate classes of G containing τ and ρ

	τ	ρ
1	1	1
Θ_i	3ϵ	$\epsilon\theta_i(\rho)$
Ξ_j	$-\epsilon$	$\epsilon\xi_j(\rho)$
E_k	$\delta\phi_k(\tau)=\pm 2u\delta$	0
X_i	δ_i	δ_i
others	0	0

TABLE II.

respectively. Then if a and b are the coefficients of $\langle\tau\rangle$ and $\langle\rho\rangle$ resp. in the expansion of $\langle\tau\rangle^2$, we have $a=2v$ and $b=0$. By Lemma 3,

$$2v=g(4uv)^{-2}\{1+\delta_1(\mathfrak{D}_gX_1)^{-1}+\delta_2(\mathfrak{D}_gX_2)^{-1}+(8\epsilon u-9\epsilon)(\mathfrak{D}_g\Theta_i)^{-1}\}$$

and

$$0=g(4uv)^{-2}\{1+\delta_1(\mathfrak{D}_gX_1)^{-1}+\delta_2(\mathfrak{D}_gX_2)^{-1} + (9\epsilon\sum_i\theta_i(\rho)+\epsilon\sum_j\xi_j(\rho))(\mathfrak{D}_g\Theta_i)^{-1}\}.$$

We used the equality $\mathfrak{D}_g\Theta_i=\mathfrak{D}_g\Xi_j$. But by the orthogonality relations we have

$$3+\sum_i\mathfrak{D}_g\theta_i\cdot\theta_i(\rho)=0, \quad\text{or}\quad \sum_i\theta_i(\rho)=-1,$$

and

$$3+\sum_i\mathfrak{D}_g\theta_i\cdot\theta_i(\rho)+\sum_j\mathfrak{D}_g\xi_j\cdot\xi_j(\rho)=0, \quad\text{or}\quad \sum_j\xi_j(\rho)=0.$$

Hence

$$0=g(4uv)^{-2}\{1+\delta_1(\mathfrak{D}_gX_1)^{-1}+\delta_2(\mathfrak{D}_gX_2)^{-1}-9\epsilon(\mathfrak{D}_g\Theta_i)^{-1}\}.$$

Together with the first relation, we obtain $g=4uv^3\epsilon\mathfrak{D}_g\Theta_i$. Since $g=4uvm$ and $(m,uv)=1$, we conclude that $v=1$, which is clearly impossible. Hence U must be the unit group and our theorem is thus proved.

As a corollary we have

THEOREM. *Let G be a finite simple group. If the order of G is $p_1^2p_2\cdots p_n$ (p_i are primes and $p_1<p_2<\cdots<p_n$) then $G\cong LF(2,p_n)$.*

5. In this section we shall prove the following Theorem B.

THEOREM B. *Let G be a non-solvable group of even order. If we assume that all l-Sylow subgroups are cyclic when $l>2$, and the 2-Sylow subgroups are of dihedral type, then G contains a normal subgroup G_1 such that $[G:G_1]\leqq 2$ and $G_1=Z\times L$, where Z is a Z-group and L is isomorphic with some $LF(2,p)$.*

Proof. We use an induction on the order of G. Throughout the proof we may assume that the order of G is divisible by 8, since the other case is covered by Theorem A. If G contains a proper non-solvable normal subgroup U, U contains a characteristic subgroup L isomorphic with some $LF(2,p)$. L is a normal subgroup of G, so that the centralizer Z of L in G is also normal, and by Lemma 2 $Z\cup L$ has index 1 or 2. The assertion of Theorem B is then proved by putting $Z\cup L=G_1$.

We assume that all proper normal subgroups of G are solvable. From this assumption we conclude in particular that G is perfect. By Theorems 5 and 6 of Zassenhaus [13], a solvable normal subgroup of G contains a normal subgroup P of G such that the order of P is either an odd prime or a power of 2. Since l-Sylow subgroups of G are cyclic and G is perfect, G does not contain any normal subgroup of order l by Theorem 15 of Zassenhaus [13]. Suppose now that G contains a normal subgroup P of order 2^{μ}. We may assume that P is a minimal one. Since 2-Sylow subgroups of G are of dihedral type, the order of P is either 2 or 4. If $[P:e]=2$, then P is in the center of G. By a theorem of Grün (see [8], or [14], p. 134), G is not perfect. (Note that S_2 is a direct factor of its normalizer, since we have assumed $[S_2:e]\geqq 8$.) If $[P:e]=4$, we consider the centralizer V of P in G. $V\neq G$, and G/V is isomorphic with a subgroup of the octahedral group, and so is solvable. In either cases we get a contradiction to our assumption of the perfectness of G. Hence G is a simple group.

Since G is perfect, elements of order 2 in G form a single conjugate class, as is seen from Grün's theorem (loc. cit.). Let τ be an element of order 2, and put $T=\{\tau\}$. Then the normalizer N of T in G contains a 2-Sylow subgroup S_2 of G and a normal subgroup H such that $HS_2=N$ and $H\cap S_2=e$. This is proved by using Grün's theorem as before.

Consider the normalizer N_2 of S_2 in G. By our assumption $[S_2:e]\geqq 8$, and so $N_2=S_2\times W$. By way of contradiction, we assume that $W\neq e$. Let q be the smallest prime divisor of $[W:e]$, and W_q a q-Sylow subgroup of W. Consider the normalizer U and the centralizer V of W_q in G. Assume that $U\neq V$. Then V is a normal subgroup of U and the factor group U/V is a cyclic group of order less than q. Since $V\supseteq S_2$, we have $U=V\bar{N}_2$, where $\bar{N}_2$ is the normalizer of S_2 in U, i.e. $\bar{N}_2=N_2\cap U$. By the second isomorphism theorem, $[U:V]=[\bar{N}_2:\bar{N}_2\cap V]$, and hence $[U:V]$ is a divisor of the order of N_2. This contradicts with the choice of q. Hence $U=V$; in other words, W_q is in the center of its normalizer. Hence by a theorem of Burnside, G is not simple. We have thus proved that $W=e$, or $N_2=S_2$.

S_2 contains a cyclic subgroup Z of index 2. Since $N_2=S_2$, the normalizer D of Z in G is a dihedral group. Hence if $N=D$, our assertion follows from Lemma 7.

By way of contradiction we assume that $N\neq D$. The subgroup H is a 2-Sylow complement of N. Put $K=H\cap D$; K is cyclic. Take an element σ of order 2 in S_2, which is not contained in Z. Then K is contained in the commutator subgroup of $\{H,\sigma\}$. This implies that K is a normal subgroup of H.

The subgroup HZ is a Z-group and hence its commutator subgroup I is cyclic. Since KZ is a maximal abelian subgroup of HZ, we have $I \cup KZ = HZ$, or $I \cup K = H$. Since we have assumed that $N \neq D$, some Sylow subgroup, say the q-Sylow subgroup I_q, of I is not contained in K. I_q is a normal subgroup of N, and Z is not contained in the centralizer of I_q, since $I_q \not\subseteq K = H \cap D$. But the centralizer of I_q is a normal subgroup of N with cyclic factor group, and hence it contains at least three elements of order 2. On the other hand, I_q is the totality of elements of N with orders which are powers of q. Hence the normalizer of I_q is larger than N. Let M be a maximal subgroup of G containing the normalizer of I_q in G. N must be a proper subgroup of M.

Suppose first that M is not solvable. Then we may apply our inductive hypothesis on M. M contains a normal subgroup M_1 such that $[M:M_1] \leqq 2$ and $M_1 = Z_1 \times L_1$, where Z_1 is a Z-group and $L_1 \cong LF(2, p')$ for some prime number p'. I_q is a subgroup of Z_1, since I_q commutes with at least three elements of order 2. Now if $M = M_1$, then I_q is in the centralizer of S_2 in G. This is impossible. Hence $[M:M_1] = 2$. Since M/Z_1 is a linear group by Lemma 2, we have $[N:Z_1] = 2(p' - \epsilon)$, where $\epsilon = (-1)^{(p'-1)/2}$. By the inductive hypothesis, the order of Z_1 is relatively prime to $p'^2 - 1$, a divisor of the order of L_1, and hence the order of N can not be divisible by $(p' + \epsilon)/2$. This gives a contradiction, since M contains an element π of order 2 such that $\pi \notin L_1$ and π commutes with an element of order $(p' + \epsilon)/2$. Hence M must be solvable.

Now by Theorem 6 of Zassenhaus [13], M contains a normal subgroup M_1 with the properties that M_1 is a Z-group and M/M_1 either has order 2 or is isomorphic with the tetrahedral or octahedral group.

If M/M_1 is of order 2, then M contains a 2-Sylow complement J. J is a normal subgroup of M and contains H as a proper subgroup. As before, K is a normal subgroup of J, since K is in the commutator subgroup of a Z-group $\{J, \sigma\}$. Similarly, I is normal in J, since I is in the commutator subgroup of JZ. Hence $H = I \cup K$ is normal in J. Since J is a Z-group, every normal subgroup of J is characteristic. Hence H is a normal subgroup of M. Consider the factor group M/H and the image $\bar{\tau}$ of the generator τ of the subgroup T of order 2 under the natural homomorphism. The normalizer of $\bar{\tau}$ in M/H is N/H, and since $H = J \cap N$, $\bar{\tau}$ induces the automorphism of J/H which maps every element into its inverse (see Lemma 1 of Zassenhaus [12]). Now N/H is isomorphic with S_2 and is generated by two elements ρ and ρ' of order 2. If we take a Sylow subgroup $\bar{P}$ of J/H, ρ (or $\rho\bar{\tau}$) commutes with all elements of P. Similarly ρ' (or $\rho'\bar{\tau}$) commutes with all elements of $\bar{P}$. Hence $\bar{P} \cup N/H = \bar{P} \times N/H$, and this is a contradiction.

The 2-Sylow subgroup S_2 coincides with its own normalizer. Hence M can not contain any proper normal subgroup with odd index. In particular M/M_1 can not be isomorphic with the tetrahedral group.

Suppose finally that M/M_1 is isomorphic with the octahedral group. Then M contains a maximal subgroup M_2 such that $[M:M_2]=2$ and M_2/M_1 is isomorphic with the tetrahedral group. A 2-Sylow subgroup of M_2 is a dihedral group and its three maximal subgroups are conjugate in M_2 to each other. Hence the order of M_2 is not divisible by 8, or M_1 is of odd order. Consider a Sylow subgroup S of M_1, and denote by X and Y the normalizer and the centralizer of S in M_2 respectively. Y is a normal subgroup of X and the factor group X/Y is cyclic, since S is cyclic. Put $Q=M_2\cap S_2$. Then Q is a non-cyclic group of order 4. Now $M_2=M_1X$, and so $X/X\cap M_1 \cong M_2/M_1$. Hence $Y\cup M_1$ is a normal subgroup of M_2 with cyclic factor group $M_2/Y\cup M_1$. Since M_2/M_1 is isomorphic with the tetrahedral group, the order of $Y\cup M_1$ is divisible by 4, and so $Y\cup M_1\supseteq Q=M_2\cap S_2$. This means that Q is contained in the centralizer of some conjugate subgroup of S. Since S was any Sylow subgroup of M_1, Q is in the centralizer of M_1, or $Q\cup M_1=Q\times M_1$, and Q is a normal subgroup of M. The factor group M/Q is of dihedral type, because S_2 is its own normalizer in G. This implies in particular that M_2/Q is cyclic. M_2 contains a subgroup R such that $RQ=M_2$ and $R\cap Q=e$. Then R is isomorphic with M_2/Q, and $R\supseteq M_1$. We want to prove that the normalizer, in G, of any subgroup of R is contained in M. Consider a subgroup X of R. If $X\cap M_1\neq e$, the normalizer of X is contained in the normalizer of $X\cap M_1$. Since M is a maximal subgroup of G and G is simple, the normalizer of $X\cap M_1$ is M, and so the normalizer of X is contained in M. If $X\cap M_1=e$, then the centralizer of X is of odd order. Hence the normalizer Y of X is a Z-group. Since R is contained in the commutator subgroup of Y, R is a normal subgroup of Y, or Y is contained in the normalizer of M_1, i.e. $Y\subseteq M$. Hence all conjugate classes of M containing elements $\neq 1$ are special with respect to G.

Consider the irreducible characters of M and M_1Q. Set $[M_1:e]=v$. Let θ_0 be the principal character of M_1Q, ω a non-principal character of M_1Q/M_1, and $\theta_1,\cdots,\theta_q$ non-principal characters of M_1Q/Q, one taken from each associated class of characters, where $q=(v-1)/2$.

If ψ is any character of M_1Q, we denote by $\bar{\psi}$ that character of M which is induced by ψ. The following relations are obtained:

$$\bar{\theta}_0=1+\eta+2\phi_0, \qquad \bar{\omega}=\xi_1+\xi_2,$$

$$\bar{\theta}_i=\phi_{3i-2}+\phi_{3i-1}+\phi_{3i}, \qquad \overline{\omega\theta_i}=\Theta_i,$$

where 1 is the principal character of M, η the non-principal linear character, $\phi_0, \cdots, \phi_{3q}$ characters of degree 2, ξ_1 and ξ_2 characters of M/M_1 of degree 3, and the Θ_i are of degree 6. All these characters of M are irreducible.

We apply Lemma 5. G has exceptional characters $\Phi_0, \cdots, \Phi_{3q}$ associated with $\phi_0, \cdots, \phi_{3q}$. If η_0 is the principal character of M_2, then $\bar{\eta}_0 = 1 + \eta$. Hence $\int_G | \eta_0^* - \phi_i^* |^2 = 3$. Since $\phi_i^* = \epsilon\Phi_i + \Delta_1$, $\eta_0^* = 1 + \cdots$ and $3q \geqq 3$,

$$\phi_i^* = \epsilon\Phi_i + a\sum\Phi_k + aX + \Delta_2 \qquad (\epsilon = \pm 1),$$

and

$$\eta_0^* = 1 + a\sum\Phi_k + (a+\epsilon)X + \Delta_2,$$

where Δ_2 does not contain X or Φ_k. (The multiplicity of X in ϕ_i^* is the same as that of Φ_k in η_0^*, since $w(\phi_i^*) = w(\eta_0^*) - 1$.) Hence

$$\theta_0^* = 1 + 2\epsilon\Phi_0 + \epsilon X + 3a\sum\Phi_k + 3aX + 3\Delta_2,$$

$$\theta_i^* = \epsilon(\Phi_{3i-2} + \Phi_{3i-1} + \Phi_{3i}) + 3a\sum\Phi_k + 3aX + 3\Delta_2.$$

Applying Lemma 4, we have

$$w((\omega\theta_i)^* - \theta_0^*) = w(\Theta_i - \bar{\theta}_0) = 7,$$

$$w((\omega\theta_i)^* - \theta_j^*) = w(\Theta_i - \bar{\theta}_j) = 4.$$

Moreover, by making computations similar to those used in the proof of Lemma 4, we obtain

$$\int_G (\omega\theta_i)^*(\bar{\Phi}_j - \bar{\Phi}_k) = \pm\int_M \Theta_i(\bar{\phi}_j - \bar{\phi}_k) = 0.$$

Suppose $(\omega\theta_i)^*$ contains Φ_k and X with the multiplicities b and c respectively. Then

$$7 = 1 + (3a-b)^2 3q + (3a+2\epsilon-b)^2 + (3a+\epsilon-c)^2 + \gamma,$$

$$4 = (3a-b)^2(3q-2) + 3(3a+\epsilon-b)^2 + (3a-c)^2 + \gamma,$$

where γ is a non-negative integer. Hence

$$6 = 2(3a-b)^2 + (3a+2\epsilon-b)^2 + (3a+\epsilon-c)^2 + \gamma',$$

$$4 = 3(3a+\epsilon-b)^2 + (3a-c)^2 + \gamma',$$

where $\gamma' = \gamma + (3a-b)^2(3q-2)$. We have two cases:

Case I. $3a - b = 0$. In this case, the above equation can be read as

$$2 = (3a+\epsilon-c)^2 + \gamma' \text{ and } 1 = (3a-c)^2 + \gamma'.$$

Hence $3a-c=0$ and $\gamma'=1=\gamma$.

Case II. $3a-b+\epsilon=0$. In this case,

$$3=(3a+\epsilon-c)^2+\gamma' \text{ and } 4=(3a-c)^2+\gamma'.$$

Hence

$$3=(3a-c)^2+2\epsilon(3a-c)+1+\gamma'=5+2\epsilon(3a-c),$$

or $3a-c=-\epsilon$ and $\gamma'=3$. Since $\gamma'=\gamma+3q-2$, we conclude that $v=3$ and $\gamma=2$.

We shall treat these cases separately.

Case I. If $v>3$, G has exceptional characters $E_1,\cdots,E_q$ associated with $\Theta_1,\cdots,\Theta_q$ (see Lemma 5). If $v=3$ we cannot apply Lemma 5; but there exists only one character E_1 of G whose multiplicity in $(\omega\theta_1)^*$ is not divisible by 3, since $\gamma=1$. We take E_1 as the exceptional character associated with Θ_1. In any case we have decompositions of induced characters in the following form:

$$\theta_0{}^*=1+2\epsilon\Phi_0+\epsilon X+3a\sum\Phi_k+3aX+3b\sum E_l+3\Delta_3,$$

$$\theta_i{}^*=\epsilon(\Phi_{3i-2}+\Phi_{3i-1}+\Phi_{3i})+3a\sum\Phi_k+3aX+3b\sum E_l+3\Delta_3 \quad (i\geqq 1),$$

$$(\omega\theta_j)^*=\epsilon E_j \qquad +3a\sum\Phi_k+3aX+3b\sum E_l+3\Delta_3 \quad (j\geqq 1).$$

Consider then ω^*; we have from Lemma 4

$$w(\omega^*-\theta_0{}^*)=8,\ w(\omega^*-\theta_i{}^*)=5 \text{ and } w(\omega^*-(\omega\theta_j)^*)=3.$$

Since $\int_G \omega^*(\Phi_i-\Phi_j)=\pm\int_M \bar{\omega}(\phi_i-\phi_j)=0$ and since

$$\int_G \omega^*(E_i-E_j)=\pm\int_M \bar{\omega}(\Theta_i-\Theta_j)=0,$$

we have the decomposition of ω^* as follows:

$$\omega^*=a'\sum\Phi_k+a''X+b'\sum E_l+\Delta_4,$$

where Δ_4 does not contain Φ_k, X and E_l.

We now get three equations for a', a'', and b', namely

$$3=(a'-3a)^2(3q+1)+(a''-3a)^2 + (b'-3b)^2(q-1)+(b'-3b-\epsilon)^2+\Gamma,$$

$$5=(a'-3a)^2(3q-2)+3(a'-3a-\epsilon)^2 + (a''-3a)^2+(b'-3b)^2q+\Gamma,$$

$$8=1+(a'-3a)^2 3q+(a'-3a-2\epsilon)^2 + (a''-3a-\epsilon)^2+(b'-3b)^2+\Gamma,$$

where $\Gamma = w(\Delta_4 - 3\Delta_3) \geqq 0$. The first two equations may be written as

$$3 = 3(a' - 3a)^2 + (b' - 3b - \epsilon)^2 + \Gamma',$$

$$5 = 3(a' - 3a - \epsilon)^2 + (b' - 3b)^2 + \Gamma' \qquad (\Gamma' \geqq 0).$$

Hence $a' = 3a$ or $a' = 3a + \epsilon$. But if $a' = 3a + \epsilon$, then the first equation shows that $b' = 3b + \epsilon$ and $\Gamma' = 0$. This solution does not satisfy the second one. Hence we have $a' = 3a$ and $b' = 3b$. If we use these values for a' and b', then the first and the last equations become

$$3 = (a'' - 3a)^2 + 1 + \Gamma, \qquad 8 = 1 + 4 + (a'' - 3a - \epsilon)^2 + \Gamma.$$

Hence $2 = (a'' - 3a)^2 - 2\epsilon(a'' - 3a) + \Gamma = 2 - 2\epsilon(a'' - 3a)$, or $a'' = 3a$, and hence $\Gamma = 2$. Now we can write

$$\omega^* = 3a \sum \Phi_k + 3aX + 3b \sum E_l + \Delta_4, \qquad \Delta_4 = \delta_1 Y + \delta_2 Z + 3\Delta_3,$$

where $\delta_i = \pm 1$ and Y and Z are irreducible characters of G, which are not Φ_k, X, E_l $(k = 0, \cdots, 3q;\ l = 1, 2, \cdots, q)$.

We can now compute the values of the characters. Let τ be an element of order 2, and ρ a generator of the subgroup R.

	1	τ	ρ
1	1	1	1
E_j	$6\epsilon + 12bv$	-2ϵ	0
Φ_i	$2\epsilon + 12av$	2ϵ	$\epsilon\phi_i$
X	$\epsilon + 12av$	ϵ	ϵ
Y	$3\delta_1 + 12yv$	$-\delta_1$	0
Z	$3\delta_2 + 12zv$	$-\delta_2$	0

TABLE III.

All other irreducible characters of G, not listed in Table III, vanish on ρ. We compute $\langle\tau\rangle^2$ in the group ring; then $\langle\rho\rangle$ has multiplicity $3v$ in $\langle\tau\rangle^2$, since the normalizer of R contains exactly $3v$ elements of order 2. Hence by Lemma 3 we have

$$3v = g(8v)^{-2}\{1 + \epsilon(\mathcal{D}_g X)^{-1} + 4\epsilon(\mathcal{D}_g \Phi_i)^{-1} \sum_i \phi_i(\rho)\}.$$

If we put $x = \mathcal{D}_g X$, then $\mathcal{D}_g \Phi_i = x + \epsilon$. An orthogonality relation yields $\sum_i \phi_i(\rho) = -1$. Hence

$$3v = g(8v)^{-2}\{1 + \epsilon x^{-1} - 4\epsilon(x+\epsilon)^{-1}\} = g(8v)^{-2}(x-\epsilon)^2 x^{-1}(x+\epsilon)^{-1},$$

or $g = 3 \cdot 8^2 v^3 (12av + \epsilon)(12av + 2\epsilon)(12av)^{-2}$, that is,

$$g = 24v \cdot (12av + \epsilon)(6av + \epsilon)/3^2a^2.$$

This is impossible, since the last fraction should be an integer.

Case II. In this case we have $v = 3$, and the induced characters are decomposed as follows:

$$\theta_0{}^* = 1 + 2\epsilon\Phi_0 + \epsilon X + 3\Delta_1, \qquad \theta_1{}^* = \epsilon(\Phi_1 + \Phi_2 + \Phi_3) + 3\Delta_1,$$

$$(\omega\theta_1)^* = \epsilon(\Phi_0 + \Phi_1 + \Phi_2 + \Phi_3) + \epsilon X + \delta_1 Y + \delta_2 Z + 3\Delta_1,$$

where $\delta_i = \pm 1$ and Y, Z are irreducible characters of G, not equal to X, Φ_i, or the principal character. In the same way as in Case I, the induced character ω^* is decomposed into a sum of irreducible characters as follows: $\omega^* = a' \sum \Phi_k + a''X + \Delta_5$, where Δ_5 does not contain Φ_k or X. If we write $\Delta_1 = a \sum \Phi_k + aX + \Delta_2$ as before, we get three equations for a' and a''. The equation $w(\omega^* - (\omega\theta_1)^*) = 3$ implies that

$$3 = 4(a' - 3a - \epsilon)^2 + a'' - 3a - \epsilon)^2 + \cdots.$$

Hence a' must be equal to $3a + \epsilon$. Similarly $w(\omega^* - \theta_0{}^*) = 8$ shows that

$$8 = 1 + 4 + (a'' - 3a - \epsilon)^2 + w(\Delta_5 - 3\Delta_2),$$

or $(a'' - 3a - \epsilon) + w(\Delta_5 - 3\Delta_2) = 3$. Finally $w(\omega^* - \theta_1{}^*) = 5$ implies that

$$5 = 1 + (a'' - 3a)^2 + w(\Delta_5 - 3\Delta_2),$$

or $(a'' - 3a)^2 + w(\Delta_5 - 3\Delta_2) = 4$. Hence $a'' = 3a + \epsilon$ and $w(\Delta_5 - 3\Delta_2) = 3$.

Let b, b' be the multiplicities of Y in Δ_2 and Δ_5 respectively, and similarly let c and c' be those of Z. Then $w(\Delta_5 - 3\Delta_2) = 3$ implies that

$$(3b - b')^2 + (3c - c')^2 + \Gamma = 3 \text{ or, since } w(\omega^* - (\omega\theta_1)^*) = 3,$$

we get $(3b + \delta_1 - b')^2 + (3c + \delta_2 + c')^2 + \Gamma = 3$. Hence $3b = b'$ or $3c = c'$. Without loss of generality we may assume that $b' = 3b$. Then $c' = 3c + \delta_2$ and $\Gamma = 2$. Hence

$$\omega^* = \epsilon \sum \Phi_k + \epsilon X + \delta_2 Z + \delta_3 V + \delta_4 W + 3\Delta_1,$$

where $\delta_3, \delta_4 = \pm 1$ and V, W are irreducible characters of G. We have the following table. (τ and ρ have the same meanings as in Table III.)

	1	τ	ρ
1	1	1	1
Φ_i	$36a + 11\epsilon$	$-\epsilon$	$\epsilon\phi_i(\rho)$
Y	$36y + 6\delta_1$	$-2\delta_1$	0
X	$36a + 10\epsilon$	-2ϵ	ϵ
Z	$36z + 9\delta_2$	$-3\delta_2$	0
V	$36v + 3\delta_3$	$-\delta_3$	0
W	$36w + 3\delta_4$	$-\delta_4$	0

TABLE IV.

All characters of G which do not appear in this table have degrees divisible by 36, and hence vanish on τ and ρ. From this table we get, by Lemma 3, the following equation:

$$9 = g(24)^{-2}\{1 + 4\epsilon(\mathfrak{D}_g X)^{-1} + \epsilon(\mathfrak{D}_g \Phi_i)^{-1} \sum_i \phi_i(\rho)\}.$$

Hence $9 = g(24)^{-2}\{1 + 4\epsilon(36a + 10\epsilon)^{-1} - \epsilon(36a + 11\epsilon)^{-1}\}$, since $\sum_i \phi_i(\rho) = -1$; so that

$$g = 9 \cdot 8 \cdot (36a + 11\epsilon)(18a + 5\epsilon)(3a + \epsilon)^{-2}.$$

It can be seen easily that $3a + \epsilon$ is relatively prime to $36a + 11\epsilon$ and $18a + 5\epsilon$. Hence $(3a + \epsilon)^2$ must be 1, i.e. $3a + \epsilon = \pm 1$, or $a = 0$. Hence $\epsilon = +1$ and $g = 8 \cdot 9 \cdot 5 \cdot 11 = 3960$.

We shall complete the first column of the table IV. Since θ_1^* and $(\omega\theta_1)^*$ have the same degree, we get

$$\epsilon\mathfrak{D}_g\Phi_i + \epsilon\mathfrak{D}_g X + \delta_1\mathfrak{D}_g Y + \delta_2\mathfrak{D}_g Z = 0,$$

or $\delta_1 y + \delta_2 z = -1$. Similarly, since $\mathfrak{D}_g(\omega\theta_1)^* = \mathfrak{D}_g\omega^*$,

$$\delta_1\mathfrak{D}_g Y = \delta_3\mathfrak{D}_g V + \delta_4\mathfrak{D}_g W,$$

or $\delta_1 y = \delta_3 v + \delta_4 w$. If one of y, z, v, w is greater than 1, then G has a representation of degree at least equal to $72 - 9 = 63$. Since $(63)^2 = 3969 > 3960 = g$, this can not happen. Hence each of the integers y, z, v, w is either 0 or 1. If $z = 1$, then y must be 0 and $\delta_2 = -1$. Hence $\mathfrak{D}_g Z = 27 = 3^3$, which is impossible. (For the order of G is not divisible by 27.) Hence $z = 0$. So we conclude that $\delta_2 = 1$, $y = 1$, and $\delta_1 = -1$, i.e., $\mathfrak{D}_g Y = 30$ and $\mathfrak{D}_g Z = 9$. Since $\delta_3 v + \delta_4 w = -1$, we may assume without loss of gnerality that $v=0$, $\delta_3=1$, $w=1$, and $\delta_4=-1$. Hence $\mathfrak{D}_g V=3$ and $\mathfrak{D}_g W=33$. The sum of the squares of the degrees of all the irreducible characters is equal to the order of G, and the degrees of the missing characters are multiples of 36. Hence

$$g=1+4(11)^2+(10)^2+(30)^2+(9)^2+(3)^2+(33)^2+k(36)^2,$$

or $k=1$. Hence there is only one character missing and the degree of the missing character is exactly 36. This is however impossible, since characters of degree 36 belong to a 2-block of defect 1, which contains exactly two characters of degree 36 (see [1]).

In both cases we have derived a contradiction, and hence the normalizer of elements of order 2 must be of dihedral type. This proves Theorem B.

6. We shall now prove the following theorem:

THEOREM C. *Let G be a non-solvable group of finite even order. We assume that all Sylow subgroups of odd order are cyclic and that a 2-Sylow subgroup is a generalized quaternion group. Then G contains a normal subgroup G_1 with the properties that $[G:G_1]\leqq 2$ and G_1 is a direct product of a Z-group and a subgroup isomorphic with $SL(2,p)$ for some prime number p.*

Proof. We shall again use the inductive argument. Assume that G is not perfect. Then the commutator subgroup G' of G is a proper subgroup satisfying the assumptions made on G. Hence we may apply the inductive hypothesis to G'. G' has the structure described in Theorem C; in particular, G' contains a characteristic subgroup T of order 2. T is clearly a normal subgroup of G. Now the factor group G/T satisfies all the assumptions of Theorem B. Hence G/T contains a normal subgroup G_1/T such that $[G:G_1]\leqq 2$ and $G_1/T\cong\bar{Z}\times\bar{L}$, where $\bar{Z}$ is a Z-group of odd order and $\bar{L}$ is isomorphic with $LF(2,p)$. Hence G_1 contains two normal subgroups L and Z such that $G_1=lZ$, $L\cap Z=T$, $\bar{L}\cong L/T$, and $\bar{Z}\cong Z/T$. Since T is normal and the 2-Sylow subgroups are of quaternion type, L contains only one subgroup of order 2. Hence by Lemma 1, L is isomorphic with $SL(2,p)$. Since $\bar{Z}$ is of odd order, Z contains a subgroup Z_1 such that $Z=Z_1\times T$. Z_1 is a normal subgroup of G_1 and $G_1=Z_1\times L$. This proves our theorem in this case.

We may, therefore, assume that G is perfect. Take a 2-Sylow subgroup S_2 of G. The center T of S_2 is of order 2. We denote by τ the generator of T and consider the normalizer N of τ in G. Since G is 2-normal in the sense of Grün [8], we have an isomorphism $G/G'(2)\cong N/N'(2)$, where

$G'(2)$ and $N'(2)$ are the 2-commutator subgroups of G and N respectively. Since we have assumed that G is perfect, N does not contain any normal subgroup of index 2.

Suppose that N is solvable. Using a theorem of Zassenhaus, we know that N contains a normal subgroup H such that N/H is isomorphic with the tetrahedral group. Hence $H \cap S_2 = T$. Now H is a Z-group, so that H contains a subgroup I such that $H = I \times T$. Let J be the union of I and S_2. Consider a Sylow subgroup X of I, the normalizer Y of X in N, and the centralizer Z of X. Then we have $N = IY$ and Z is a normal subgroup of Y with cyclic factor group Y/Z. Hence ZI is a normal subgroup of N with a cyclic factor group. This implies that the index $[N:ZI]$ is odd, and hence S_2 is in the centralizer of some conjugate subgroup of X. Since X was any Sylow subgroup of I, we conclude that $J = I \cup S_2 = I \times S_2$. Hence J must be the centralizer of S_2 in G. Now N contains a subgroup K such that $KS_2 = N$, $K \cap J = I$, and $[K:I] = 3$. If the order of I were divisible by 3, then the normalizer and the centralizer of a 3-Sylow subgroup would coincide with each other, and so G would not be perfect. Hence the order of I is not a multiple of 3. Consider then a subgroup P of K with order 3. P is contained in a 3-Sylow subgroup S_3 of G. Let N_3 and Z_3 be the normalizer and the centralizer of S_3 in G respectively. Since G is perfect, Z_3 is a normal subgroup of N_3 with index 2. If $[Z_3:e]$ were odd, there would be an element σ of order 2 such that σ induces an automorphism of S_3 which maps every element into its inverse. Hence the normalizer of P in G would contain two elements of order 2, σ and τ, such that one of them commutes with every element in P but the other does not. This is impossible, since σ and τ are conjugate in the normalizer of P. Hence the order of Z_3 is even. This implies that G contains an element ν of order 4 such that $\nu\pi\nu^{-1} = \pi^{-1}$ for every $\pi \varepsilon P$. The normalizer of P in G contains ν, and ν^2 is conjugate to τ in this normalizer. This is clearly not the case, because of the structure of N. Hence N is not solvable.

The factor group N/T satisfies all the assumptions of Theorem B, but does not contain any normal subgroup of index 2. Hence as before N is a direct product of a Z-group Z and a linear group. If $Z \neq e$, we consider a Sylow subgroup X of Z which belongs to the smallest prime divisor of the order of Z. The normalizer Y of X in G is a non-solvable proper subgroup of G. Hence we may apply the inductive hypothesis and conclude that T is in the center of Y; in other words, Y is contained in N. From the choice of X, we see that Y coincides with the centralizer of X. Hence by a theorem of Burnside, G is not perfect. Hence we have $Z = e$, or N is isomorphic with

$SL(2,p)$ for some p. Now our Theorem C is a consequence of the following general proposition.

THEOREM D. *Let G be a group of finite even order. If G contains an element τ of order 2 such that the normalizer of τ in G is isomorphic with the special linear group $SL(2,p)$ for some $p > 2$, then G contains an abelian normal subgroup G_0 such that $G/G_0 \cong SL(2,p)$, and if $G_0 \neq e$, then every primary component of G_0 is not cyclic.*

Proof. Denote by N the normalizer of the element τ of order 2. By assumption N is isomorphic with $SL(2,p)$. In the following we consider only the case $p \equiv 1 \pmod 4$. The other case will be treated in exactly the same way, with a slight modification on signs.

Put $p-1=2^\alpha h$, $(h,2)=1$. Then $\alpha \geqq 2$ and $(p+1)/2 = 2a+1$ is odd. We shall consider the structure of 2-blocks of G with positive defect. By Theorem 4 of Brauer [2], G has exactly a 2-blocks of defect 1, $(h-1)/2$ blocks of defect α, and one block of defect $\alpha+1$. We have a one-to-one correspondence between blocks of positive defect of G and those of N. Namely if we take a block B of positive defect d $(d>0)$ of G, then B determines a block B of defect d in the normalizer $N(D)$ of the defect group D of B. Since $d>0$, D is not the unit group, and we may take D as a subgroup of N. Then T, the subgroup generated by τ, is a characteristic subgroup of D. Hence the normalizer $N(D)$ is also a subgroup of N. Hence $\tilde{B}$ determines a block B^* of N. In this manner we get a one-to-one correspondence $B \leftrightarrow B^*$. We want to show that B and B^* consist of exactly the same number of irreducible characters. Consider any subgroup X of N, whose order is a power of 2, and a block $\bar{B}$ of the normalizer Y of X in G. Then $\bar{B}$ determines a block B of G and a block B' of N. Let $\bar{\varpi}$ be the linear character of the center of the modular group ring of Y, corresponding to the block $\bar{B}$. Then the corresponding linear character ω for B satisfies

$$\omega(C) \equiv \sum \bar{\varpi}(C_\lambda) \pmod{\mathfrak{p}}$$

where $\mathfrak{p}$ is a prime divisor of 2 in a suitable algebraic number field and C_λ ranges over all classes of Y which lie in C and whose elements belong to the centralizer of Y. Similarly the linear character ω' of B' is defined by

$$\omega'(C') \equiv \sum \bar{\varpi}(C'_\mu) \pmod{\mathfrak{p}}\text{; hence } \omega(C) \equiv \sum \omega'(C'_\nu) \pmod{\mathfrak{p}},$$

where C'_ν ranges over all classes of N which belong to C. This implies that $B' = B^*$ in the previous notation. Hence if a block $\bar{B}$ of X determines a block B of G, then $\bar{B}$ determines a block B^* of N. We apply Theorems 2 and 3 of Brauer [3] to G and N, and conclude that the number of characters in B is equal to that in B^*.

Let B_1 be the first 2-block of N. Then $\bar{B}_1$ consists of the principal character 1, one character ϕ_1 of degree p, two conjugate characters ϕ_2 and ϕ_3 of degree $(p+1)/2$, two characters ϕ_4 and ϕ_5 of degree $(p-1)/2$, and $2^{\alpha-1}-1$ characters ϕ_i $(i=6,7,\cdot\cdot\cdot,2^\alpha+4)$ of degree $p+1$ (see Section 1 of this paper). Let ϕ be a linear combination of these characters which vanishes on all 2-regular elements, say $\phi = 1+\phi_1-\phi_2-\phi_3$. Since all 2-singular classes of N are special with respect to G, we may apply Lemma 4. It follows that $w(\phi^*)=w(\phi)=4$. On the other hand, ϕ^* contains the principal character of G with multiplicity 1. Hence

$$\phi^* = 1+\epsilon_1 X_1+\epsilon_2 X_2+\epsilon_3 X_3 \qquad (\epsilon_i=\pm 1).$$

Similarly we get

$$(1-\phi_2+\phi_4)^* = 1+\delta_1 Y_1+\delta_2 Y_2 \qquad (\delta_i=\pm 1).$$

But $(1+\epsilon_1 X_1+\epsilon_2 X_2+\epsilon_3 X_3)-(1+\delta_1 Y_1+\delta_2 Y_2) = (\phi_1-\phi_3+\phi_4)^*$ is a sum of three irreducible characters. Hence

$$(1-\phi_2+\phi_4)^* = 1+\epsilon_2 X_2+\epsilon_4 X_4 \qquad (\epsilon_4=\pm 1).$$

Similarly

$$(1-\phi_3+\phi_5)^* = 1+\epsilon_3 X_3+\epsilon_5 X_5 \qquad (\epsilon_5=\pm 1),$$

and

$$(\phi_i-\phi_2-\phi_3)^* = \epsilon_i X_i+\epsilon_2 X_2+\epsilon_3 X_3 \qquad (\epsilon_i=\pm 1).$$

These characters X_i are all different. Any one of the above relations between the X_i vanishes on all 2-regular classes. Hence by Theorem 6 of Brauer-Nesbitt [4], these relations remain true if we omit all characters which are not in the fixed 2-block B of G. Hence if some X_k is not in the first 2-block B_1 of G, then X_k must vanish on all 2-regular classes. This is clearly impossible. Hence all the X_k are in B_1 and B_1 contains no other characters. Using Lemma 4, we obtain the following table (for notations see Table I):

	2-regular	order 2	order 2^α	$p+1$
$1+\epsilon_1 X_1+\epsilon_2 X_2+\epsilon_3 X_3$	0	0	4	0
$1+\epsilon_2 X_2+\epsilon_4 X_4$	0	$1-p$	2	2
$1+\epsilon_3 X_3+\epsilon_5 X_5$	0	$1-p$	2	2
$\epsilon_i X_i+\epsilon_2 X_2+\epsilon_3 X_3$	0	$\begin{cases}0\\-2(p+1)\end{cases}$	$2+\theta_i$	0

TABLE V.

These relations are linear equations of the values of X_k for elements of order 2, etc. We want to solve these equations. Before doing so, we shall derive other relations.

If $\bar{B}_\kappa$ is a block of defect α of N, $\bar{B}_\kappa$ consists of 2^α characters $\theta_1, \cdots, \theta_{2^\alpha}$ of degree $p+1$. They take the same value on every 2-regular class. We apply Lemma 5. The exceptional characters $\Theta_1, \cdots, \Theta_{2^\alpha}$ associated with these $\theta_1, \cdots, \theta_{2^\alpha}$ belong to the same block B_κ of G, since they are identical on 2-regular classes. This block has defect greater than 2, since $\alpha \geqq 2$. In the first 2-block B_1, we have several degree relations, for instance $1 + \epsilon_1 \mathfrak{D}_g X_1 - \epsilon_i \mathfrak{D}_g X_i = 0$. Hence at least one of X_1 and X_i is not exceptional. Hence B_1 contains at least 4 characters which are not exceptional for θ_i. Hence B_κ is a block of defect α. Since each block of defect α contains exactly 2^α characters, B_κ contains only exceptional characters $\Theta_1, \cdots, \Theta_{2^\alpha}$. By definition of exceptional characters, Θ_i is decomposed as $\Theta_i = \epsilon\theta_i + \Delta$ $(i = 1, 2, \cdots, 2^\alpha)$ on N. Hence if σ is any 2-singular element of G, we have $\Theta_i(\sigma) = \epsilon\theta_i(\sigma) + c$, where c is independent of i. The orthogonality relation yields

$$\sum_i (\mathfrak{D}_g \Theta_i)\Theta_i(\sigma) = 0, \text{ or } \sum \Theta_i(\sigma) = 0.$$

Similarly $\sum_i \theta_i(\sigma) = 0$. Hence we have $c2^\alpha = 0$, $c = 0$. Hence we conclude that $\Theta_i(\sigma) = \epsilon\theta_i(\sigma)$ for any 2-singular element. If we take another block $\bar{B}_\lambda$ of defect α, then the exceptional characters of G associated with characters in $\bar{B}_\lambda$ form a block B_λ of defect α, and $B_\lambda \neq B_\kappa$. Hence every block of defect α of G is obtained in the above process of taking exceptional characters.

If B is a block of defect 1 of G, B consists of two characters Ψ_1 and Ψ_2 which take the same value on every 2-regular element. Hence we have

$$2 = w(\Psi_1 - \Psi_2) = \int_N |\Psi_1 - \Psi_2|^2,$$

or $\Psi_1 - \Psi_2 = \psi_1 - \psi_2$ on N, where ψ_1 and ψ_2 are irreducible characters of N. ψ_i is not in any block of defect α. For if ψ_1 belongs to a block $\bar{B}$ of defect α, then ψ_2 will belong to $\bar{B}$. Since $(\psi_1 - \psi_2)^* = \Psi_1 - \Psi_2$, Ψ_i must be exceptional characters associated with ψ_i. But ψ_i belongs to $\bar{B}$, and so Ψ_i belongs to a block B of defect α. This is not the case. Hence ψ_i does not belong to any block of defect α. Since

$$\int_G (\Psi_1 - \Psi_2)(1 + \epsilon_2 X_2 + \epsilon_4 X_4) = \int_N (\psi_1 - \psi_2)(1 - \phi_2 + \phi_4) = 0,$$

ψ_i is not ϕ_2, and similarly ψ_i is not equal to any $\phi_i \varepsilon \bar{B}_1$. Hence ψ_1 and ψ_2 are in a block of defect 1. Hence for any 2-singular element of G, we have $\Psi_i(\sigma) = \psi_i(\sigma)$ $(i = 1, 2)$.

Let σ be any 2-singular element of G. The orthogonality relation yields $\sum_{\mu} | \chi_{\mu}(\sigma) |^2 = n(\sigma)$, where χ_{μ} ranges over all irreducible characters of G and $n(\sigma)$ is the order of the normalizer of σ. Since the normalizer of σ is contained in N, we have the same relation for N, namely, $\sum_{\nu} | \phi_{\nu}(\sigma) |^2 = n(\sigma)$, where ϕ_{ν} ranges over all irreducible charatcers of N. Now any character of defect 0 vanishes on σ. If B is not the first 2-block of G, then the contribution of the characters in B to $\sum_{\mu}$ is equal to that of the corresponding block of N to $\sum_{\nu}$. Hence we conclude that

$$(*) \qquad \sum_{i} | X_i(\sigma) |^2 = \sum_{i} | \phi_i(\sigma) |^2$$

where the summation is extended over $i = 1, 2, \cdots, 2^{\alpha-1} + 5$. We apply this relation for $\sigma = \tau, \rho, v$, where τ is an element of order 2, ρ an element of order 2^{α} and v of order $p + 1$.

Put $x_i = \epsilon_i X_i(\tau)$. Then we have from table V

$$\begin{aligned} 1 + x_1 + x_2 + x_3 &= 0 \\ 1 + x_2 + x_4 &= 1 - p \\ 1 + x_3 + x_5 &= 1 - p \\ x_i + x_2 + x_3 &= 0 \qquad \text{or } -2(p+1). \end{aligned}$$

Put $x_2 = s$ and $x_3 = t$. Then s and t are real. The relation (*) is now written as

$$\begin{aligned} 1 + (1 + s + t)^2 + s^2 + t^2 + (p + s)^2 + (p + t)^2 + (s + t)^2(2^{\alpha-2} - 1) \\ + (s + t + 2(p + 1)^2(2^{\alpha-2}) \\ = 1 + p^2 + \tfrac{1}{2}(p - 1)^2 + \tfrac{1}{2}(p + 1)^2 + (p + 1)^2(2^{\alpha-1} - 1), \end{aligned}$$

or $(2^{\alpha-1} + 1)(p + 1 + s + t)^2 + (s - t)^2 = 0$. Since s and t are real, we have $s = t = -\frac{1}{2}(p + 1)$. Now we obtain the first column of the following table VI.

	τ	ρ	ν
1	1	1	1
x_1	$\epsilon_1 p$	ϵ_1	$-\epsilon_1$
x_2	$-\epsilon_2 \frac{1}{2}(p + 1)$	ϵ_2	0
x_3	$-\epsilon_2 \frac{1}{2}(p + 1)$	ϵ_3	0
x_4	$-\epsilon_4 \frac{1}{2}(p - 1)$	0	ϵ_4
x_5	$-\epsilon_5 \frac{1}{2}(p - 1)$	0	ϵ_5
x_i	$\begin{cases} \epsilon_i(p + 1) \\ -\epsilon_i(p + 1) \end{cases}$	$\epsilon_i \theta_i$	0

TABLE VI.

Similarly the relation (*) for $\sigma = \rho$ is written as

$$1 + (3 - u - v)(3 - \bar{u} - \bar{v}) + u\bar{u} + v\bar{v} + (1 - u)(1 - \bar{u}) + (1 - v)(1 - \bar{v})$$
$$+ \sum_i (\theta_i + 2 - u - v)(\theta_i + 2 - \bar{u} - \bar{v}) = 4 + \sum_i \theta_i \bar{\theta}_i$$

where $u = \epsilon_2 X_2(\rho)$, $v = \epsilon_3 X_3(\rho)$ and the bar indicates the conjugate complex numbers. Using the relation $\sum_i \theta_i = 0$, we get

$$(2^{\alpha-1} + 1)(u + v - 2)(\bar{u} + \bar{v} - 2) + (u - v)(\bar{u} - \bar{v}) = 0,$$

or $u = v = 1$. Hence we obtain the second column of table VI.

For the last column, we get

$$1 + (1 + w + z)(1 + \bar{w} + \bar{z}) + w\bar{w} + z\bar{z} + (1 - w)(1 - \bar{w})$$
$$+ (1 - z)(1 - \bar{z}) + (2^{\alpha-1} + 1)(w + z)(\bar{w} + \bar{z}) = 4,$$

setting $w = \epsilon_2 X_2(\nu)$ and $z = \epsilon_3 X_3(\nu)$. Hence

$$2^{\alpha-1}(w + z)(\bar{w} + \bar{z}) + 2w\bar{w} + 2z\bar{z} = 0,$$

or $w = z = 0$. This completes the last column of table VI.

Since N is isomorphic with $SL(2, p)$, N contains only one element of order 2. Hence there is no element of order 2 which transforms ρ or ν into its inverse. This means that neither ρ nor ν appears in the expansion $\langle\tau\rangle^2$. Hence by Lemma 3 we obtain two relations:

$$1 + p^2\epsilon_1(\mathcal{D}_g X_1)^{-1} + (\tfrac{1}{2}(p + 1))^2\epsilon_2(\mathcal{D}_g X_2)^{-1} + (\tfrac{1}{2}(p + 1))^2\epsilon_3(\mathcal{D}_g X_3)^{-1} = 0,$$

and

$$1 - p^2\epsilon_1(\mathcal{D}_g X_1)^{-1} + (\tfrac{1}{2}(p - 1))^2\epsilon_4(\mathcal{D}_g X_4)^{-1} + (\tfrac{1}{2}(p - 1))^2\epsilon_5(\mathcal{D}_g X_5)^{-1} = 0.$$

Note that characters belonging to a block B of defect 1 or α do not contribute to these sums, since $\sum_{s \in B} s(\sigma) = 0$ for all 2-singular elements σ. We obtain moreover the following degree relations from Table V.

$$1 + \epsilon_1\mathcal{D}_g X_1 + \epsilon_2\mathcal{D}_g X_2 + \epsilon_3\mathcal{D}_g X_3 = 0$$
$$1 + \epsilon_2\mathcal{D}_g X_2 + \epsilon_4\mathcal{D}_g X_4 = 0$$
$$1 + \epsilon_3\mathcal{D}_g X_3 + \epsilon_5\mathcal{D}_g X_5 = 0.$$

We can solve these equations for $\epsilon_i\mathcal{D}_g X_i$. Put $\xi_i = \epsilon_i\mathcal{D}_g X_i$ $(i = 1, 2, \cdots, 5)$. Then the first two relations are

$$\xi_1\xi_2\xi_3 + p^2\xi_2\xi_3 + [\tfrac{1}{2}(p + 1)]^2\xi_1(\xi_2 + \xi_3) = 0,$$
$$\xi_1\xi_4\xi_5 - p^2\xi_4\xi_5 + [\tfrac{1}{2}(p - 1)]^2\xi_1(\xi_4 + \xi_5) = 0.$$

We use the relations $-\xi_4 = 1 + \xi_2$, $-\xi_5 = 1 + \xi_3$ and make these substitutions in the second equation. Then we get

$$\xi_1(1+\xi_2)(1+\xi_3) - p^2(1+\xi_2)(1+\xi_3) - [\tfrac{1}{2}(p-1)]^2\xi_1(2+\xi_2+\xi_3) = 0.$$

Now put $X = \xi_1$ and $Y = \xi_2\xi_3$; then $\xi_2 + \xi_3 = -(1+X)$. We have then two equations for X and Y, i.e.

$$XY + p^2Y - [\tfrac{1}{2}(p+1)]^2X(1+X) = 0,$$

$$X(Y-X) - p^2(Y-X) - [\tfrac{1}{2}(p-1)]^2(1-X)X = 0.$$

Add these two equations and then divide by X. We get

$$2Y = X(p+1) - \tfrac{1}{2}(p^2-1).$$

Substituting this value in the first equation, we obtain

$$(X+p^2)\tfrac{1}{2}(p+1)^2(X-\tfrac{1}{2}(p-1)) - [\tfrac{1}{2}(p+1)]^2(1+X)X = 0,$$

or $X^2 - 2pX + p^2 = 0$, and hence $X = p$.

Since $X = \xi_1 = \epsilon_1 \mathfrak{D}_g X_1$, we must have $\epsilon_1 = 1$ and $\mathfrak{D}_g X_1 = p$. From table VI, we see that $X_1(\tau) = p$. This means that τ is contained in the kernel K of the representation corresponding to X_1. K is a normal subgroup of G, and hence $K \cap N$ is a normal subgroup of N, which contains τ but not ρ. Since $N \cong SL(2, p)$, the only normal subgroup of N which contains τ but not ρ is the center $T = \{\tau\}$. Hence $K \cap N = T$. By assumption N is the normalizer of τ in G. Hence the normalizer of τ in K coincides with T. By a theorem of Burnside, K contains a normal subgroup G_0 such that $G_0T = K$ and $G_0 \cap T = e$. If σ is any element $\neq 1$ of G_0, σ does not commute with τ, so G_0 admits an automorphism of order 2 which leaves only the identity invariant. By a lemma of Zassenhaus G_0 must be abelian.

We have to prove that the factor group G/G_0 is isomorphic with $SL(2, p)$. $K = G_0T$ is a normal subgroup of G, and T is a 2-Sylow subgroup of K, since G_0 is of odd order. Hence G is a join of K and the normalizer of T in G. But N is the normalizer of T, so that $G = KN$, or $G = G_0N$. Hence $G/G_0 \cong N/G_0 \cap N = N$.

Let P be any Sylow subgroup of G_0. If P is cyclic and $\neq e$, then the centralizer Q of P in G is a normal subgroup with cyclic factor group. $Q \cap N$ is then a normal subgroup of N, but it does not contain T. Hence $Q \cap N = e$; but this is impossible because N is not cyclic. Hence every Sylow subgroup of G_0 is not cyclic. This completes the proof of Theorem D for the case where $p \equiv 1 \pmod 4$.

If $p \equiv 3 \pmod 4$, we can use the same method by changing the signs suitably. Hence our Theorem D, and therefore Theorem C also, is proved.

It can be proved that Theorem D is still true if we assume that the normalizer of an element of order 2 is isomorphic with $SL(2,q)$ over any finite field with prime power $q=p^{\mu}$ elements $(p>2)$. The proof is quite similar to the proof given here, and depends upon those properties of irreducible characters of $SL(2,q)$ which are given by Jordan [9] and Schur [10].

7. We shall add a few remarks. It is well-known that if every abelian subgroup of a finite group G is cyclic, then Sylow subgroups of G are either cyclic or a generalized quaternion group. If moreover G is non-solvable, then a 2-Sylow subgroup must be a generalized quaternion group. Thus our Theorem C gives the structure of such groups. The complete characterization may be found in Theorem E below.

There are many occasions that we have a finite group G in which every abelian subgroup is cyclic. We shall mention here two cases: (a) G is a multiplicative group contained in some non-commutative field, and (b) the cohomology groups of G over the rational integers are periodic. In both cases every abelian subgroup of G is cyclic. (For the latter case (b) cf. the recent work by Artin and Tate.)

Finally we shall consider the converse to our main theorem. Let G be a finite non-solvable group, all of whose abelian subgroups are cyclic. Then by Theorem C, G contains a normal subgroup G_1 such that $[G:G_1] \leqq 2$ and $G_1 = Z \times L$, where Z is a Z-group and $L \cong SL(2,p)$. But a group having this structure may contain an abelian subgroup which is not cyclic. For instance, the group consisting of all 2×2 matrices over the prime field with determinant $+1$ or -1 has the above structure, i.e. it contains $SL(2,p)$ as a subgroup of index 2, but it contains non-cyclic subgroups of order 4. If G_1 coincides with G, then a necessary and sufficient condition is simply that the order of Z is relatively prime to that of L. However if G_1 is a proper subgroup of G, the condition is not so simple. To state the condition explicitly we define an automorphism θ of $SL(2,p)$ by

$$\theta:\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}\rightarrow\begin{pmatrix}1 & 0\\ -\omega & 1\end{pmatrix},\quad \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\rightarrow\begin{pmatrix}0 & \omega^{-1}\\ -\omega & 0\end{pmatrix}$$

where ω is a primitive element, i.e. a generator of the multiplicative group of the underlying field. In fact, θ is induced in $SL(2,p)$ by the matrix transformation by $\begin{pmatrix}0 & -1\\ \omega & 0\end{pmatrix}$. It is easy to see that $\theta^2=1$ and θ is not an inner automorphism of $SL(2,p)$. Moreover if f is any automorphism of $SL(2,p)$ which is not inner, then f induces an automorphism $\bar{f}$ of the central factor group $LF(2,p)$. If $\bar{f}$ were inner, then $SL(2,p)$ would admit a non-trivial automorphism of the form $\lambda \rightarrow \lambda\tau(\lambda)$ where $\tau(\lambda)$ is a central element. Since $\tau(\lambda)$ would be a homomorphism of $SL(2,p)$ into its center, $\tau(\lambda)$ would be trivial. Hence $\bar{f}$ is not inner. By Lemma 2, $\theta^{-1}f$ induces an inner automorphism of the central factor group. Using the same argument as above, we conclude that $\theta^{-1}f$ is an inner automorphism of $SL(2,p)$. This θ has the following property.

LEMMA 9. *If $\lambda\theta(\lambda)$ is in the center of L for some $\lambda \varepsilon L$, then $\lambda\theta(\lambda)=1$, where $L \cong SL(2,p)$.*

Proof. We may use the matrix notation. Let $\lambda = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$, $\alpha\delta - \beta\gamma = 1$. Then $\lambda\theta(\lambda)$ is in the center if and only if

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}\begin{pmatrix} 0 & -1 \\ \omega & 0 \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}\begin{pmatrix} 0 & \omega^{-1} \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon \end{pmatrix},$$

where $\epsilon = \pm 1$. A little computation shows that $\epsilon = 1$.

Using this θ, we now state the final theorem.

THEOREM E. *Every abelian subgroup of a non-solvable group G of finite order is cyclic if and only if either* (a) *G is a direct product of a Z-group Z and a group isomorphic with $SL(2,p)$ for some $p > 3$ and having order relatively prime to the order of Z, or* (b) *G is generated by $L \cong SL(2,p)$ and three elements, σ, ρ and τ such that*

$$\sigma^m = \rho^n = \tau^4 = 1,$$

$$\rho\sigma\rho^{-1} = \sigma^r, \qquad \tau\sigma\tau^{-1} = \sigma^{-1}, \qquad \tau\rho = \rho\tau$$

$$\rho\lambda = \lambda\rho, \qquad \sigma\lambda = \lambda\sigma, \qquad \tau\lambda\tau^{-1} = \theta(\lambda)$$

for any $\lambda \varepsilon L$, and τ^2 is in the center of L. Here m, n, and r are integers such that

$$(m,n) = 1, \quad (mn, p(p^2-1)) = 1 \text{ and } r^n \equiv 1 \pmod{m}.$$

Proof. Suppose that every abelian subgroup of G is cyclic. Then G contains a subgroup G_1 such that $[G:G_1] \leqq 2$ and $G_1 = Z \times L$ where Z is a Z-group and $L \cong SL(2,p)$. Since all abelian subgroups are cyclic, the order of Z must be prime to $p(p^2-1)$. If $G_1 = G$, then we are finished.

Assume $G \neq G_1$. Then G contains a subgroup L_0 such that L is a subgroup of L_0 with index 2. Since the 2-Sylow subgroups of L_0 are generalized quaternion groups, L_0 contains an element π such that $\pi \notin L$ and π^2 is in the center of L. π induces an automorphism of L, which is not inner. Hence there exists an element $\lambda_0 \varepsilon L$ such that $\pi\lambda\pi^{-1} = \lambda_0\theta(\lambda)\lambda_0^{-1}$ for every $\lambda \varepsilon L$. Since π^2 is in the center of L, we must have

$$\lambda = \pi^2\lambda\pi^{-2} = \lambda_0\theta(\lambda_0)\lambda(\lambda_0\theta(\lambda_0))^{-1},$$

or $\lambda_0\theta(\lambda_0)$ is an element of the center of L. By Lemma 9, we get $\lambda_0\theta(\lambda_0) = 1$. Let $\tau = \lambda_0^{-1}\pi$. Then

$$\begin{aligned}\tau^2 &= \lambda_0^{-1}\pi\lambda_0^{-1}\pi = \lambda_0^{-1}(\pi\lambda_0^{-1}\pi^{-1})\pi^2 \\ &= \lambda_0^{-1}(\lambda_0\theta(\lambda_0)^{-1}\lambda_0^{-1})\pi^2 = (\lambda_0\theta(\lambda_0))^{-1}\pi^2 = \pi^2.\end{aligned}$$

Since G/L is a Z-group, we can take a set of generators $\{\sigma, \rho\}$ of Z which satisfies all the requirements of Theorem E.

Conversely we assume that G has the structure in this theorem. Then it is easy to see that any l-Sylow subgroup of G is cyclic if $l > 2$. It remains to show that a 2-Sylow subgroup is a generalized quaternion group. It suffices now to show that a subgroup $\{L, \tau\}$ contains only one element of order 2. Suppose $(\lambda\tau)^2 = 1$ for some $\lambda \varepsilon L$. Then $1 = \lambda\tau\lambda\tau = \lambda(\tau\lambda\tau^{-1})\tau^2 = \lambda\theta(\lambda)\tau^2$. It follows therefore that $\lambda\theta(\lambda)$ is in the center. By Lemma 9, this implies that $\lambda\theta(\lambda) = 1$, or $\tau^2 = 1$. This is clearly impossible. Hence the 2-Sylow subgroups of G are generalized quaternion groups and our assertion is proved.

UNIVERSITY OF ILLINOIS.

REFERENCES.

[1] R. Brauer, "Investigations on group characters," *Annals of Mathematics*, vol. 42 (1941), pp. 936-958.

[2] ———, "On the arithmetic in a group ring," *Proceedings of the National Academy of Science*, vol. 30 (1944), pp. 109-114.

[3] ———, "On blocks of characters of groups of finite order. I," *Proceedings of the National Academy of Science*, vol. 32 (1946), pp. 182-186.

[4] ——— and C. Nesbitt, "On the modular characters of groups," *Annals of Mathematics*, vol. 42 (1941), pp. 556-590.

[5] ——— and K. A. Fowler (to appear).

[6] ———, M. Suzuki and H. S. Wall (to appear).

[7] W. Burnside, *The Theory of Groups*, ed. 1911.

[8] O. Grün, "Beiträge zur Gruppentheorie. I," *Journal für die reine und angewandte Mathematik*, vol. 174 (1936), pp. 1-14.

[9] H. E. Jordan, "Group-characters of various types of linear groups," *American Journal of Mathematics*, vol. 29 (1907), pp. 387-405.

[10] I. Schur, "Untersuchungen über die Darstellung der endlichen Gruppen durch gebrochene lineare Substitutionen," *Journal für die reine und angewandte Mathematik*, vol. 132 (1907), pp. 85-137.

[11] M. Suzuki, "A characterization of simple groups $LF(2,p)$," *Journal of the University of Tokyo*, vol. VI-4 (1951), pp. 259-293.

[12] H. Zassenhaus, "Kennzeichung endlicher linearer Gruppen als Permutationsgruppen," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 11 (1936), pp. 17-40.

[13] ———, "Über endliche Fastkörper," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 11 (1936), pp. 187-220.

[14] ———, *Lehrbuch der Gruppentheorie*, vol. I, 1937.

ASYMPTOTIC INTEGRATIONS OF ORDINARY NON-LINEAR DIFFERENTIAL EQUATIONS.*

By PHILIP HARTMAN and AUREL WINTNER.

Introduction. It is well-known that Poincaré's classical results on systems of differential equations $dy^i/dt = f^i(y^1, \cdots, y^d)$, $i = 1, \cdots, d$, where the functions f^i are power series convergent near, and vanishing at, the origin and have a non-vanishing Jacobian there, can be extended so as to dispose of the conditions of analyticity; in fact, this matter has an extensive literature. In what follows, the problem will be taken up in its full generality, by methods previously applied by us in particular cases (for the non-conservative linear cases, cf. [6]; for binary non-linear cases, cf., e. g., [3], [4]). The generality of the methods will have two different aspects. On the one hand, nothing like a Lipschitz condition will be needed (so that no local uniqueness is assumed) and, on the other hand, the matrix of the linear approximation is allowed to have characteristic numbers and elementary divisors of arbitrary multiplicities. The continuous functions f^i will be allowed to depend on t also.

As in [6], the problem of the asymptotic integration of the system near $(y^1, \cdots, y^d) = (0, \cdots, 0)$ will be treated at two different stages of refinement; first, on a "logarithmic scale" (Part I) and, second, on the basis of the resulting rough asymptotic formulae, without this logarithmic reduction (Part II). The results of the available literature on the first stage are more complete than those on the second stage.

In order to describe the problems in a more detailed form, consider the system of differential equations

$$(1) \qquad y' = Jy + F(t, y)$$

(with $y' = dy/dt$), where J is a constant d by d matrix, y and F are d-dimensional (column) vectors, and $F(t; y)$ is defined and "small" for large t and small $|y|$. The method to be employed is a combination of the procedure involving Riccati differential equations as majorants (in the same way as in the linear cases treated in [2], [6]) and of the procedure involving "zero velocity curves" (as in the binary cases of (1) treated in [3]). The variant of the "zero velocity curve" argument (cf. Sections 7, 13 below)

---

* Received January 19, 1955.

replaces that involving the superposition principle in Sections 7 and 16 of [6].

For the sake of simplicity, only the case where the elements of J, F and y are complex-valued will be treated. Hence there is no loss of generality in supposing that J is in its Jordan normal form, say, that the linear system

$$y' = Jy \tag{2}$$

has the form

$$y^{j1\prime} = \lambda y^{j1}, \tag{3_{j1}}$$

$$y^{jk\prime} = \lambda y^{jk} + y^{j\,k-1}, \qquad k = 2, \cdots, h, \tag{3_{jk}}$$

where $\lambda = \lambda(j)$; $h = h(j)$; $j = 1, \cdots, g$; $h(1) + \cdots + h(g) = d$; and (3_{jk}), for $k > 1$, is missing if $h(j) = 1$. Corresponding to the notation y^{jk}, where $j = 1, \cdots, g$ and $k = 1, \cdots, h(j)$, for the d components of y, let $F^{jk} = F^{jk}(t, y)$ denote the corresponding d components of $F(t, y)$. Thus (1) has the form

$$y^{j1\prime} = \lambda y^{j1} + F^{j1}(t, y), \tag{4_{j1}}$$

$$y^{jk\prime} = \lambda y^{jk} + y^{j\,k-1} + F^{jk}(t, y), \qquad k = 2, \cdots, h. \tag{4_{jk}}$$

It will be clear from the proofs that the methods to be used supply analogous theorems for the case in which the elements of J, F and y are real-valued; so that the normal form of J involves binary blocks of the type

$$\begin{pmatrix} \mu & \nu \\ -\nu & \mu \end{pmatrix}.$$

Part I of the paper deals with the asymptotic integration of (1) on a logarithmic scale. Theorems (*), (**) below are analogues of Theorem (*) of [6] in the linear cases of (1) and of Satz 11 of Perron [11], pp. 148-149 (cf. [12]), dealing with non-linear systems. In Theorems (*) and (**) below, the condition on $F(t, y)$ is that $|F(t, y)| / |y| \to 0$ as $(t, y) \to (\infty, 0)$ (or even relaxed forms of this condition; cf. Section 10). Perron [11], p. 149, requires a Lipschitz condition of the type

$$|F(t, y_1) - F(t, y_2)| / |y_1 - y_2| \to 0 \text{ as } (t, y_1, y_2) \to (\infty, 0, 0),$$

and his assertion is less sharp than the assertions (11), (12) of (*), (**) below. On the other hand, Perron's theorem contains an assertion of uniqueness. Such an assertion in (*) is impossible, since the conditions on $F(t, y)$ do not even imply the local uniqueness of solutions of (1). (See Section 2, for references to known cases of (*), (**), with $d = 2$.)

Part II deals with asymptotic formulae for solutions $y = y(t)$ of (1), rather than with asymptotic formulae for the logarithms of their magnitudes. The theorems (†), (††) below are analogues of the refinement (**) in [6] of Dunkel's theorem [1] dealing with linear systems. The theorems (†), (††) are known in cases of conservative binary ($d=2$) systems. For references to the literature, see Section 12 below.

Part I. Solutions tending to the origin.

1. Notation. Let $\lambda(1),\cdots,\lambda(g)$ denote the eigenvalues of J belonging to its g elementary divisors; cf. (3). It is not assumed that the numbers $\lambda(1),\cdots,\lambda(g)$ are distinct. Let $\mu(j)=\operatorname{Re}\lambda(j)$ and let $\mu^1<\cdots<\mu^f$ denote the distinct numbers among the g numbers $\mu(1),\cdots,\mu(g)$. For a fixed m, where $1\leqq m\leqq f$, an integer j in the set $1,2,\cdots,g$ will be denoted by p, q or r according as $\mu(j)<\mu^m$, $\mu(j)=\mu^m$ or $\mu(j)>\mu^m$:

$$(5)\qquad \mu(j)\lesseqgtr\mu^m \text{ according as } j=p,\ q \text{ or } r.$$

Put

$$(6)\qquad L_m=\sum_q\sum_{k=1}^{h(q)}|y^{qk}|^2;$$

so that $|y|^2=L_1+\cdots+L_f$ is the square of the length of the vector y.

From time to time, it will be convenient to use the conventional notation $y=(y^1,\cdots,y^d)$ for the components of y and to consider the set, say $\lambda_1,\cdots,\lambda_d$, of the d eigenvalues of J. In such cases, it will be assumed that the numeration is chosen so that the k-th component y^k corresponds to the k-th eigenvalue λ_k, in the sense that the differential equation for y^k in (1) or (2) is of the form $y^{k\prime}=\lambda_k y^k+\cdots$.

2. Theorems (*) and (**). In terms of the above notations, the following theorem will first be proved:

(*) *Let the linear system of differential equations* (2) *with* (*complex-valued*) *constant coefficients have the form* (3). *Let* $F(t,y)$ *be a vector* (*with complex-valued components*) *defined and continuous for* $0\leqq t<\infty$, $|y|\leqq\text{const.}$ *and let*

$$(7)\qquad |F(t,y)|/|y|\to 0 \text{ as } (t,y)\to(\infty,0).$$

Let the number $\mu=\mu^m$ *be negative,*

$$(8)\qquad \mu^m<0.$$

Then, if η *is any number on the range* $0 < \eta < 1$, *there exist positive numbers* $\delta = \delta(\eta)$ *and* $T = T(\eta)$ *with the property that if* $t_0 \geqq T$ *and if the set of* $\Sigma h(p) + \Sigma h(q)$ *numbers* y_0^{pk} *and* y_0^{qk} *satisfies*

$$(9_1) \quad \sum_p \sum_k |y_0^{pk}|^2 \leqq \eta \sum_q \sum_k |y_0^{qk}|^2, \qquad (9_2) \quad 0 < \sum_q \sum_k |y_0^{qk}|^2 < \delta,$$

the system (1) *has a solution* $y = y(t)$ *which exists on the half-line* $t_0 \leqq t < \infty$, *satisfies the* (*partial set of*) *initial conditions*

$$(10_1) \quad y^{pk}(t_0) = y_0^{kp}, \qquad (10_2) \quad y^{qk}(t_0) = y_0^{qk},$$

and, as $t \to \infty$, *the limit relations*

$$(11) \quad L_j = o(L_m) \text{ if } j \neq m, \qquad (12) \quad \log |y(t)| = (\mu^m + o(1))t.$$

It is understood that if $m = 1$, so that $\mu^m = \min(\mu(1), \cdots, \mu(g))$ and the set of integers p is vacuous, then (9_1) and (10_1) are missing. If $m = f$, so that $\mu^m = \max(\mu(1), \cdots, \mu(g))$ and the set of integers r is vacuous, then (10_1)-(10_2) constitutes a complete set of initial conditions.

The cases of (*) in which it is assumed that $h(q) = 1$ (that is, that the elementary divisors belonging to the $\lambda(k)$ satisfying $\operatorname{Re} \lambda(k) = \mu^m$ are simple) are contained in [15], pp. 293-294, and [13], pp. 25-26, under the additional assumption of local uniqueness.

A partial converse of (*) is as follows:

(**) *If* $F(t, y)$ *is a continuous vector* (*with complex-valued components*) *for* $0 \leqq t < \infty$, $|y| \leqq$ const. *satisfying* (7) *and if* $y = y(t)$ $(\not\equiv 0)$ *is a solution of* (1) *for large* t *such that*

$$(13) \qquad y(t) \to 0 \text{ as } t \to \infty,$$

then there is an integer m, $1 \leqq m \leqq f$, *with the properties that* $\mu = \mu^m$ *satisfies* $\mu \leqq 0$ *and that* (11) *and* (12) *hold.*

In order to illustrate the meaning of the assertion (11), suppose that $y = (y^1, y^2)$ is a binary vector, that $\mu_1 \neq \mu_2$, where $\mu_k = \operatorname{Re} \lambda_k$, and that $\mu_1 < 0$. Then, by (*), the system (1) has solutions $y(t) = (y^1(t), y^2(t))$ satisfying (13) and

$$(14) \qquad y^2(t)/y^1(t) \to 0 \text{ as } t \to \infty.$$

According to (**), if $\mu_2 > 0$, then all solutions $y = y(t)$ $(\not\equiv 0)$ of (1) satisfying (13) also satisfy (14). But if $(\mu_1 \neq)$ $\mu_2 \leqq 0$, then all solutions $y(t)$ $(\not\equiv 0)$ satisfying (13) satisfy either (14) or

$$(15) \qquad y^1(t)/y^2(t) \to 0 \text{ as } t \to \infty.$$

No positive assertion of the type (14) or (15) is possible when $\mu_1 = \mu_2$ (< 0), without further restrictions on F; cf. [5], p. 499. The conservative binary cases of (*) and (**) are contained in [7], [8], [3], [4].

3. A lemma. Lemmas 1 bis and 2 bis of Section 3 in [6], dealing with Riccati differential inequalities, will be needed below. In addition, there will be needed the following O-variant of Lemma 1 bis:

LEMMA. *Let* $c > 0$, $0 < c_0 < \frac{1}{2}$ *and let* c^0 *denote a number satisfying*

$$0 < c^0 < c_0 \text{ and } c^0 < \min cv(1-v) \text{ for } c_0 \leqq v \leqq 2c_0. \tag{16}$$

Let $f = f(t)$ *be a real-valued, continuous function on* $T \leqq t < \infty$ *satisfying*

$$f(t) \leqq c^0, \qquad (T \leqq t < \infty), \tag{17}$$

or, more generally,

$$\int_U^{U+V} f(t)\,dt \leqq c^0(1 + V) \text{ for } T \leqq U < U + V. \tag{17 bis}$$

Finally, let $v = v(t)$ *be a real-valued, continuously differentiable function on some interval* $(T \leqq)$ $t_0 \leqq t \leqq t_1$, *satisfying*

$$v' \geqq cv(1-v) - f(t) \tag{18}$$

and

$$v(t_0) \geqq 2c_0. \tag{19}$$

Then, for $t_0 \leqq t \leqq t_1$,

$$v(t) > c_0. \tag{20}$$

In the case (17), the inequality (20) can trivially be improved to $v(t) \geqq 2c_0$ since $v'(t) > 0$ whenever $v(t) = 2c_0$.

In order to prove the assertion under the assumption (17 bis), suppose that (20) fails to hold for some t, where $t_0 \leqq t \leqq t_1$. Then $[t_0, t_1]$ contains an interval $(U, U + V)$ such that $v(U) = 2c_0$, $v(U + V) = c_0$ and $c_0 \leqq v(t) \leqq 2c_0$ on $(U, U + V)$. A quadrature of (18) and the second part of (16) give

$$v(U + V) - v(U) > c^0 V - \int_U^{U+V} f(t)\,dt.$$

The left side of this inequality is $c_0 - 2c_0 = -c_0 < -c^0$. Hence

$$\int_U^{U+V} f(t)\,dt > c^0(1+V).$$

Since this contradicts (17 bis), the Lemma follows.

4. Proof of (*). Preliminaries. Let the number a, where $0 < a < 1$, be chosen so small that

$$c = c_m = 2(\mu^m - \mu^{m-1} - 2a) > 0 \quad \text{for} \quad m = 2, \cdots, f. \tag{21}$$

Put

$$L = L_{ma} = \sum_q \sum_{k=1}^{h(q)} a^{2k} \, | y^{qk} |^2, \tag{22}$$

so that

$$a^{-2} L_{ma} \leqq L_m \leqq a^{-2d} L_{ma}. \tag{23}$$

For $1 \leqq m \leqq f$, put

$$M = M_m = \sum_{j=1}^{m-1} L_{ja} \quad \text{and} \quad N = N_m = \sum_{j=m}^{f} L_{ja}. \tag{24}$$

Then, for a solution $y = y(t)$ of (1) on some interval, the expressions (22), (24) become functions of t. The function (22) then satisfies

$$| L_{ma}' - 2\mu^m L_{ma} | \leqq 2a L_{ma} + d \, | F(t, y) | \cdot | y |, \tag{25_m}$$

and so

$$M' \leqq 2(\mu^{m-1} + a) M + d^2 \, | F(t, y) | \cdot | y |, \tag{26_m}$$

$$N' \geqq 2(\mu^m - a) N - d^2 \, | F(t, y) | \cdot | y |; \tag{27_m}$$

cf. Section 4 in [6].

Although the conditions of (*) on (1) do not imply the unique determination of solutions of (1) by their initial conditions, any solution $y = y(t)$ vanishing at some large value of t, say $t \geqq T_0$, is identically 0. This is clear from condition (7) of (*). Hence, if $y = y(t) \not\equiv 0$ is a solution of (1) on some interval $U \leqq t \leqq U + V$, where $U \geqq T_0$, then $y(t) \neq 0$. Hence $M + N \geqq a^{2d} | y |^2 > 0$. Consequently, it is possible to define an auxiliary function $v = v_m(t)$ by

$$v = N_m / (N_m + M_m), \tag{28}$$

and this function satisfies

$$0 \leqq v \leqq 1 \tag{29}$$

and

$$v' \geqq c_m v(1-v) - d^2a^{-2d}\,|F(t,y)|/|y|; \tag{30}$$

cf. Section 4 in [6].

If $m < f$, then it is possible to define the function $w = w_m(t)$ $(= v_{m+1}(t))$ by

$$w = N_{m+1}/(N_m + M_m), \tag{31}$$

and this function satisfies

$$0 \leqq w \leqq 1 \tag{32}$$

and has a continuous derivative satisfying

$$w' \geqq c_{m+1}w(1-w) - d^2a^{-2d}\,|F(t,y)|/|y|. \tag{33}$$

5. The case $m = f$. As in the linear case treated in [6], the case $m = f$ is quite simple. If $m = f$, the real parts μ^j of all characteristic numbers of J satisfy $\mu^j \leqq \mu < 0$. This will be shown to imply that if $a > 0$ is fixed so that $\mu + 2a < 0$, then there exist positive numbers $T = T(a)$ and $\epsilon = \epsilon(a)$ with the property that if

$$\sum_p \sum_k |y_0^{pk}|^2 + \sum_q \sum_k |y_0^{qk}|^2 < \tfrac{1}{2}\epsilon a^{2d} \tag{34}$$

and if $t_0 \geqq T$, then a solution satisfying the (complete set of) initial conditions (10_1)-(10_2) has extensions over the half-line $t_0 \leqq t < \infty$ and all such extensions satisfy

$$|y| < \epsilon \text{ for } t \geqq t_0. \tag{35}$$

In order to see this, let $T = T(a)$ and $\epsilon(a)$ be chosen so that (35) and

$$t \geqq T \tag{36}$$

imply

$$d^2a^{-2d}\,|F(t,y)| \leqq 2a\,|y|. \tag{37}$$

The existence of such T and ϵ is clear from (7).

Since $m = f$, it follows from (24) that $|y|^2 \leqq a^{-2d}M_{m+1}$. Thus if $y = y(t)$ satisfies (35) on some t-interval $t_0 \leqq t \leqq t^0$, where $t_0 \geqq T$, then (26_{m+1}) and (37) show that

$$M_{m+1}' \leqq 2(\mu + 2a)M_{m+1}. \tag{38}$$

Hence, $M_{m+1}(t) \leqq M_{m+1}(t_0)e^{2(\mu+2a)(t-t_0)} \leqq M_{m+1}(t_0)$ on this interval, and so

$$|y(t)|^2 \leqq a^{-2d}\,|y(t_0)|^2 e^{2(\mu+2a)(t-t_0)}. \tag{39}$$

Let (34) hold. Consider a solution $y = y(t)$ of (1) determined by the initial conditions (10_1)-(10_2). Then, by continuity, (35) holds for t near $t_0\ (\geqq T)$. By (34) and (39),

$$|y(t)|^2 \leqq \tfrac{1}{2}\epsilon e^{(\mu+2a)(t-t_0)} \leqq \tfrac{1}{2}\epsilon \tag{40}$$

for such $t\ (\geqq t_0)$. Consequently, any extension of $y(t)$, with increasing t, continues to satisfy (35) (in fact, (40)). Hence $y = y(t)$ has extensions over $t_0 \leqq t < \infty$ and all extensions satisfy (40). This proves the assertion concerning (34) and (35).

Let δ in (9_2) be so small that $2\delta < \frac{1}{2}\epsilon a^{2d}$, where $\epsilon = \epsilon(a)$ is the number defined in connection with (35), (36). Since $\eta < 1$, the inequality (34) follows from (9_1) and (9_2). Hence any solution $y = y(t)$ of (1) satisfying (10_1)-(10_2) has extensions over $t_0 \leqq t < \infty$ and any extension satisfies (40). Consequently,

$$|F(t, y(t))|/|y(t)| \to 0 \text{ as } t \to \infty. \tag{41}$$

Since $\eta < 1$, condition (9_1) implies that

$$v(t_0) \geqq 2c_0, \text{ where } c_0 = 1/2(1+\eta) < \tfrac{1}{2}. \tag{42}$$

Let $T = T(a, \eta)$ be so large that (35) and (36) imply

$$d^2 a^{-2d} |F(t, y)| \leqq c^0 |y|, \tag{43}$$

if c^0 is a constant satisfying (16), where c is c_m and c_0 is given in (42). Hence, by (30), (42) and the Lemma of Section 3, (20) holds for $t_0 \leqq t < \infty$. By (29), (41) and Lemma 1 bis in Section 3 of [6],

$$v(t) \to 1 \text{ as } t \to \infty. \tag{44}$$

The definition (28) of $v(t)$ and the limit relation (44) imply that, as $t \to \infty$,

$$L_{ja}(t) = o(L_{ma}t) \text{ if } j \neq m. \tag{45}$$

By virtue of (23), this proves assertion (11) of (*).

Since, by (45), $|y|^2 \leqq a^{-2d}(1 + o(1))L_{ma}$, the limit relation (41) shows that the differential inequality (25_m) can be written as

$$|(\log L_{ma})' - 2\mu^m| \leqq 2a + o(1).$$

Hence,

$$\limsup_{t\to\infty} |t^{-1}\log L_{ma} - 2\mu^m| \leqq 2a, \text{ and so } \limsup_{t\to\infty} |t^{-1}\log L_m - 2\mu^m| \leqq 2a,$$

by (23). Since $a > 0$ can be chosen arbitrarily small, (12) follows. This completes the proof of (*) if $m = f$.

6. The case $m < f$. If a solution $y = y(t)$ of (1) on some t-interval has an auxiliary function $v = v(t)$ satisfying (20) and an auxiliary function $w = w(t)$ satisfying

$$(46) \qquad w(t) < \tfrac{1}{2}c_0$$

for some constant $c_0 > 0$, then the definitions (28), (31) of v, w show that

$$(47) \qquad L_{ma} > \tfrac{1}{2}c_0(M_m + L_{ma} + N_{m+1});$$

and since $|y|^2 \leqq a^{-2d}(M_m + L_{ma} + N_{m+1})$,

$$(48) \qquad |y|^2 \leqq 2a^{-2d}c_0^{-1}L_{ma}.$$

Thus, when (20) and (46) hold, the differential inequality (25_m) shows that

$$(49) \qquad |L_{ma}' - 2\mu L_{ma}| \leqq \{2a + 2da^{-2d}c_0^{-1}|F(t,y)|/|y|\}L_{ma}.$$

Let $a > 0$ be so small that $\mu + 2a < 0$ and let $T = T(a, \eta)$ and $\epsilon = \epsilon(a, \eta) > 0$ be chosen so that (35) and (36) imply the inequalities

$$(50) \qquad 2da^{-2d}c_0^{-1}|F(t,y)| \leqq 2a|y|$$

and (43), where, in the latter, c^0 is a constant satisfying (16) in which c is c_m or c_{m+1} and c_0 is $1/2(1+\eta)$.

Thus, if (20), (46) and

$$(51) \qquad 2a^{-2d}c_0^{-1}L_{ma} < \epsilon$$

hold, then

$$(52) \qquad L_{ma}' \leqq 2(\mu + 2a)L_{ma}, \qquad (\mu + 2a < 0).$$

On any interval, $t_0 \leqq t \leqq t^0$, on which (52) holds,

$$(53) \qquad L_{ma}(t) \leqq L_{ma}(t_0)e^{2(\mu+2a)(t-t_0)} \leqq L_{ma}(t_0).$$

Consider a solution $y = y(t) \not\equiv 0$ on some t-interval, bounded by $t = t_0$ $(\geqq T)$ on the left, satisfying (42), (46) and (51) at $t = t_0$. For an extension of $y(t)$, with increasing t, there are two possibilities: either $y(t)$ exists on $t_0 \leqq t < \infty$ and satisfies (20), (46) and (51) throughout or there exists a least value of $t = t^0$ $(> t_0)$ such that $y(t)$ satisfies (20) and (51) on $t_0 \leqq t \leqq t^0$, and (46) on $t_0 \leqq t < t^0$, finally

$$(54) \qquad w(t) = \tfrac{1}{2}c_0$$

at $t = t^0$.

In order to see this, note that if (20), (46) and (51) hold on $t_0 \leqq t < t^0$, so that (35) holds by (48), then $y(t)$ has extensions over an interval containing $t_0 \leqq t < t^0$ in its interior. Thus, the assertion above is merely to the effect that if (20), (46) and (51) hold on $t_0 \leqq t < t^0$, then (20), (51) also hold at $t = t^0$. In fact, (20), (46) and (51) imply (53) for $t_0 \leqq t < t^0$, so that (51) holds at t^0 since it holds at t_0. In view of (30), (36) and (43), the Lemma of Section 3 implies that (20) also holds on $t_0 \leqq t \leqq t^0$.

Note that the differential inequality (33) and the inequalities (20) (hence (48)) and (51) imply

$$w'(t) > 0 \text{ when } w(t) = \tfrac{1}{2}c_0 \text{ and } t \geqq T; \tag{55}$$

cf. (33) and (43).

7. Smooth $F(t, y)$. The proof of (*) will now be completed for the case that the solutions of (1) are uniquely determined (locally) by initial conditions (say, that $F(t, y)$ has continuous first derivatives with respect to $y^1, \cdots, y^d$). The proof will depend on a generalization of the "zero velocity curve" argument used for binary cases in [3]; this generalization can be considered as a particular case of the topological argument formulated in [14].

It will be shown that *to the partial set of conditions* (10_1)-(10_2), *where* $\delta > 0$ *is sufficiently small in* (9_2), *there belongs at least one set of* $\Sigma h(r)$ *numbers* y_0^{rk} *such that the solution determined by* (10_1), (10_2) *and*

$$y^{rk}(t_0) = y_0^{rk} \tag{56}$$

exists for $t_0 \leqq t < \infty$ *and satisfies* (46), *that is,*

$$N_{m+1} < \tfrac{1}{2}c_0(M_m + L_{ma} + N_{m+1}), \tag{57}$$

In other words, it will be shown that, among the set of $\Sigma h(r)$ numbers y_0^{rk} satisfying

$$(1 - \tfrac{1}{2}c_0) \sum_r \sum_k | a^k y_0^{rk} |^2 \leqq \tfrac{1}{2}c_0 (\sum_p \sum_k | a^k y_0^{pk} |^2 + \sum_q \sum_k | a^k y_0^{qk} |^2), \tag{58}$$

there is at least one set such that the solution of (1) determined by (10_1)-(10_2) and (56) exists for $t_0 \leqq t < \infty$ and satisfies (57) on this half-line.

For any choice of (56), condition (9_1) implies (36), hence (20). If (58) holds, then $w(t) \leqq \tfrac{1}{2}c_0$, and so (48) holds at $t = t_0$. Let δ in (9_2) be so small that (9_2) implies (51) at $t = t_0$.

Consider the (unique) solution determined by (10_1)-(10_2) and (56), when (58) holds. Corresponding to each choice of (56), define a number t^0 $(\leqq \infty)$ as follows: If equality holds in (58), let $t^0 = t_0$; if inequality holds in (58), let t^0 $(> t_0)$ be the least t-value $(\leqq \infty)$ at which (54) holds. According to Section 6, $y(t)$ has an extension on such an interval $t_0 \leqq t < t^0$

$(\leqq \infty)$ and (20), (46) and (51) hold on $t_0 \leqq t < t^0$. Furthermore, if $t^0 < \infty$, then (20), (51) and (54) are satisfied and (55) is applicable at $t = t^0$.

If $t^0 = \infty$ for some choice of (56) satisfying (58), then the italicized statement concerning (56) is proved. Suppose, if possible, that $(t_0 \leqq)\ t^0 < \infty$ for every choice of (56). Clearly, $y(t^0) \neq 0$. The mapping $(t_0, y(t_0)) \rightarrow (t^0, y(t^0))$ is one-to-one, by the local uniqueness of solutions of (1), and is continuous, by the continuous dependence of solutions of (1) on initial conditions, and by the fact that (55) is applicable at $t = t^0$ (so that $y = y(t)$ intersects, and is not tangent to, the "conical" surface $N_{m+1} - \frac{1}{2} c_0 (N_m + L_{ma} + N_{m+1}) = 0$ at $t = t^0$).

The domain of definition, say D, of the mapping $(t_0, y(t_0)) \rightarrow (t^0, y(t^0))$ is the set of points with the $d + 1$ coordinates $(t_0; y_0^{pk}, y_0^{qk}, y_0^{rk})$, where t_0, y_0^{pk}, y_0^{qk} are fixed and (58) holds. Since this is a closed set, and since $t^0 < \infty$ and $y(t^0) \neq 0$, there exist positive constants such that $t_0 \leqq t^0 \leqq \text{const.}$ and $|y(t^0)| \geqq \text{Const.} > 0$. Hence, the range, say R, of the mapping is contained in the product set of the t-interval $t_0 \leqq t \leqq \text{const.}$ and the "lateral surface of the frustrum of a cone":

$$(59) \quad \{t_0 \leqq t \leqq \text{const.}\} \times \{(1 - \tfrac{1}{2} c_0) N_{m+1} = \tfrac{1}{2} c_0 (M_m + L_{ma}),\ 0 < c_1 \leqq M_m + L_{ma} \leqq c_2\},$$

where c_1 and c_2 are constants. Topologically, (59) can be considered as a product set $I_1 \times I_2 \times C(d_1) \times C(d_2)$, where I_1 and I_2 are closed (real, linear) intervals, $C(j)$ is the boundary of a j-dimensional sphere (that is, of a $2j$-dimensional real sphere), and $d_1 = \Sigma h(p) + \Sigma h(q)$, $d_2 = \Sigma h(r)$. Such a representation results, for example, by writing a point of the set (59) as $(t, \rho, e(d_1), e(d_2))$, where $t_0 \leqq t \leqq \text{const.}$, $c_1^{\frac{1}{2}} \leqq \rho \leqq c_2^{\frac{1}{2}}$, $(a^k y^{pk}, a^k y^{qk}) = \rho e(d_1)$, $(1 - \frac{1}{2} c_0)(a^k y^{rk}) = \frac{1}{2} c_0 \rho e(d_2)$, and $e(j)$ is an arbitrary unit vector with j (complex-valued) components.

Since the definition of t^0 includes the identity $t^0 = t_0$ when equality holds in (58), the set $C = C(d_2)$ of points $(t; y^{pk}, y^{qk}, y^{rk})$, where $t = t_0$, $y^{pk} = y_0^{pk}$, $y^{qk} = y_0^{qk}$ and equality holds in (58), is both in the domain D and in the range R of the mapping $(t_0, y(t_0)) \rightarrow (y^0, y(t^0))$. Consider a deformation of C and D into the point $(t_0, y_0^{pk}, y_0^{qk}, 0)$. The topological mapping $(t_0, y(t_0)) \rightarrow (t^0, y(t^0))$ induces a deformation of C in R (hence in (59)) into a point. In view of the product set structure of (59), a projection leads to a deformation of C in itself into a point. Since this is impossible, the assumption that $t^0 < \infty$ holds for all points of D is untenable. This proves the italicized statement concerning (56).

The proof of (*) for the case at hand can now be completed at once.

Let $y=y(t)$ be a solution determined by (10_1), (10_2) and (56), where (56) is chosen so that $y(t)$ exists on $t_0 \leqq t < \infty$ and satisfies (46) there. Then (20) (and (51)) holds for $t_0 \leqq t < \infty$. In particular, (35), and so (41), holds.

In view of (20), (30) and (41), Lemma 1 bis in Section 3 of [6] implies (44). Similarly, it follows from (33), (46) and (41) that

$$(60) \qquad w(t) \to 0 \text{ as } t \to \infty.$$

The definitions (22), (28), (31) of L_{ma}, v, w show that (44) and (60) are equivalent to (45). The passage from (45) to (11) and (12) is the same as in the case $m=f$ treated in Section 5 above.

This proves (*) provided that the solutions of (1) are uniquely determined (locally) by initial conditions.

8. Completion of the proof of (*). In order to prove (*) without this proviso, note that if t^0 $(> T)$ is fixed, then there exists a sequence of (vector) functions $F_1(t,y)$, $F_2(t,y)$, $\cdots$ which are defined and smooth (say, of class C^1) for $T \leqq t \leqq t^0$, $|y| \leqq$ const. and such that

$$(61) \qquad F_n(t,y) \to F(t,y) \text{ as } n \to \infty,$$

uniformly for $T \leqq t \leqq t^0$, $|y| \leqq$ const. It can also be supposed that if y and t $(\leqq t^0)$ satisfy (35) and (36) respectively, then (43_n) and (50_n) hold, where (43_n) and (50_n) result from (43) and (50), respectively, by writing F_n in place of F. It can then be supposed that the domain of definition of $F_n(t,y)$ has been extended over $T \leqq t < \infty$, $|y| \leqq$ const., so that $F_n(t,y)$ is smooth and that (43_n) and (50_n) hold whenever (35) holds (whether or not $t \leqq t^0$).

After letting $t^0 = T+1, T+2, \cdots$ and applying a diagonal process, it can be supposed that the sequence $F_1(t,y). F_2(t,y), \cdots$ is independent of t^0. Thus (61) holds uniformly on every set $T \leqq t \leqq$ Const., $|y| \leqq$ const., for arbitrary Const. $(> T)$. It can also be supposed that (35) and (36) imply (43_n) and (50_n) for *every n*.

According to the considerations of Section 7, the system of differential equations

$$(62) \qquad y' = Jy + F_n(t,y)$$

has a solution $y=y_n(t)$ on $(T \leqq)\ t_0 \leqq t < \infty$ satisfying (10_1)-(10_2), (20), (46) and (51). By virtue of (46), the functions M_m, L_{ma}, N_{m+1} belonging to $y_n(t)$ satisfy $(1-\frac{1}{2}c_0)N_{m+1} < \frac{1}{2}c_0(L_{ma}+M_m)$. Since $\frac{1}{2}c_0 < 1$ and since L_{ma},

M_m are independent of n at $t=t_0$, by (10_1)-(10_2), it follows that the components $y^{rk}(t_0)=y_n^{rk}(t_0)$ of $y_n(t_0)$ are bounded (as functions of n). Hence there exists a subsequence of the integers $1,2,\cdots$ (which can be supposed to be the full sequence after a renumeration) such that

$$\lim_{n\to\infty} y_n(t)=y(t) \text{ exists} \tag{63}$$

at $t=t_0$. (In view of (10_1)-(10_2), this merely requires the existence of the limits

$$\lim_{n\to\infty} y_n^{rk}(t_0)=y_0^{rk}$$

for every (rk).)

The uniformity of the limit process (61) implies that, after another selection, it can be supposed that (63) holds uniformly on every bounded t-interval $t_0\leqq t\leqq$ Const. and that the limit $y=y(t)$ is a solution of (1) on $t_0\leqq t<\infty$.

The solution $y=y(t)$ of (1) satisfies (10_1)-(10_2) since $y=y_n(t)$ does. In view of (63), $y(t)$ satisfies the inequalities which result from (20), (46) and (51) if $<$ is replaced by $\leqq$. Finally, it can be verified as in Section 5 that $y=y(t)$ satisfies (11) and (12). This completes the proof of (*).

9. Proof of (**). The function $v(t)=v_m(t)$ belonging to the solution $y=y(t)$ in the statement of (**) satisfies (29) and (30) for large t, where (7) and (13) imply (41). Lemma 1 bis in Section 3 of [6] shows that

$$(64_1)\quad \lim_{t\to\infty} v_m(t)=1 \qquad \text{or} \qquad (64_2)\quad \lim_{t\to\infty} v_m(t)=0$$

for $m=1,2,\cdots,f$. It is clear from the definitions (24) and (28), that if (64_2) holds for some $m=m_0$, then (64_2) holds for $m\geqq m_0$. Also, (64_1) holds if $m=1$, since $v_1(t)\equiv 1$. Let m denote the largest integer (with $1\leqq m\leqq f$) for which (64_1) holds.

Thus $v(t)=v_m(t)$ satisfies (44) and, if $m<f$, $w(t)=v_{m+1}(t)$ satisfies (60). These relations imply (45) and, therefore, (11) and (12); cf. the end of Section 5. Finally, $\mu\leqq 0$ follows from (12) and (13). This proves (**).

10. Variants of (*). It will be shown that the condition (7) in (*) can be replaced by an "averaged" form of (7), a condition of the type employed in [2] and in the linear case [6] of (1). The methods of proof of this fact will have applications below.

(* bis) *Assertion* (*) *remains valid if* (9_2) *is replaced by*

$$e^{2(\mu-\delta)t}<\sum_q\sum_k |y_0^{qk}|^2<e^{2(\mu+\delta)t} \tag{65}$$

and if condition (7) *on* $F(t,y)$ *is replaced by the requirement that, for sufficiently small* $\delta > 0$, *there exists a continuous function* $f(t) = f(t;\delta)$ *with the properties that*

$$|F(t,y)| \leqq |y| f(t) \text{ if } e^{(\mu-\delta)t} < |y| < e^{(\mu+\delta)t} \tag{66}$$

and

$$\underset{0 \leqq V < \infty}{\text{l.u.b.}} \int_{U}^{U+V} f(t)\,dt/(1+V) \to 0 \text{ as } U \to \infty. \tag{67}$$

The changes in the proof of (*) needed to prove (* bis) will be indicated. It will be seen that difficulties arise from the fact that (55) is no longer valid, but that these difficulties can be avoided.

As above, let a denote a small positive number. In view of (66) and (67), there exists a $T = T(a, \eta)$ such that

$$e^{(\mu-2a)t} < |y| < e^{(\mu+2a)t} \tag{68}$$

implies the inequality in (66),

$$2d^2a^{-2d}c_0^{-1} \int_{U}^{U+V} f(t)\,dt \leqq 2a + 2aV \text{ if } T \leqq U \leqq U+V \tag{69}$$

and

$$d^2a^{-2d} \int_{U}^{U+V} f(t)\,dt \leqq c^0(1+V) \text{ if } T \leqq U \leqq U+V, \tag{70}$$

where c^0 satisfies the inequalities in (16) for $c = c_m$ and $c = c_{m+1}$ and for $c_0 = 1/2(1+\eta)$, as in (42), or if c_0 is replaced by $c_0/4 = 1/8(1+\eta)$.

Note that the applications of the Lemma in the proof of (*) remain valid if (43) is replaced by (70) (that is, if (17) is replaced by (17 bis)).

For brevity, consider the proof of (* bis) only in the case $m < f$. If (20) and (46) hold, then (47) is valid and (48) can be changed to

$$L_{ma} \leqq |y|^2 \leqq 2a^{-2d}c_0^{-1}L_{ma}. \tag{71}$$

The inequality (49) can be written in the form

$$|L_{ma}' - 2\mu L_{ma}| \leqq \{2a + 2da^{-2d}c_0^{-1}f(t)\}L_{ma} \tag{72}$$

if $t \geqq T$ and (67) hold. In view of (70), the relations (68) hold if

$$e^{2(\mu-2a)t} < L_{ma} < \tfrac{1}{2}a^{2d}c_0^{-1}e^{2(\mu-2a)t}. \tag{73}$$

When (73), hence (72), is valid on an interval $t_0 \leqq t < t^0$, a quadrature and (69) give

$$e^{-2a} e^{2(\mu-2a)(t-t_0)} \leqq L_{ma}(t)/L_{ma}(t_0) \leqq e^{2a} e^{2(\mu+2a)(t-t_0)}. \tag{74}$$

Hence (73) is assured at $t = t^0$ if

$$e^{2a} e^{2(\mu-2a)t} < L_{ma}(t) < e^{-2a} \left(\tfrac{1}{2} a^{2d} c_0{}^{-1} e^{2(\mu-2a)t}\right) \tag{75}$$

holds at $t = t_0$.

Thus, if $y = y(t)$ satisfies (42), (46) and (75) at $t = t_0$ ($\geqq T$), then (20) and (73) hold on $t_0 \leqq t \leqq t^0$ and (46) holds on $t_0 \leqq t < t^0$.

The condition (75) at $t = t_0$ is implied by (10_1)-(10_2) and the inequalities (9_1) and (65) if $\delta > 0$ is sufficiently small.

It is clear that *if* (55) *holds*, then the remainder of the proof of (*) becomes applicable to (* bis) (if the inequalities (9_2) and (51) in that proof are replaced by (65) and (73) respectively). It will be seen that the possible failure of (55) is not a serious difficulty.

Let $F_0(t, y)$ be a continuous function on the set of (t, y) satisfying inequalities $T \leqq t < \infty$, (68) and

$$N_{m+1} \leqq \tfrac{1}{2} c_0 (M_m + L_{ma} + N_{m+1}). \tag{76}$$

Let F_0 have the properties that $F_0(t, y) \equiv F(t, y)$ when

$$N_{m+1} \leqq (c_0/4)(M_m + L_{ma} + N_{m+1}); \tag{77}$$

that, for other points (t, y) of the domain of definition of F_0,

$$|F_0(t, y)| \leqq |F(t, y)| + \theta, \tag{78}$$

where $\theta = \theta(a)$ is a positive number to be described below; finally, that

$$d^2 a^{-2d} |F_0(t, y)| \leqq c^0 |y| \quad \text{if} \quad N_{m+1} = \tfrac{1}{2} c_0 (M_m + L_{ma} + N_{m+1}). \tag{79}$$

In the considerations involving (68)-(75) and the definition of F_0, let a be replaced by a smaller positive number, say $\frac{1}{2}a$, and, correspondingly, $T = T(a, \eta)$ by $T(\frac{1}{2}a, \eta)$. Then, if $\theta = \theta(a) > 0$ in (78) is sufficiently small, it is clear that the remarks concerning (68)-(75) remain valid if $y = y(t)$ is a solution of

$$y' = Jy + F_0(t, y). \tag{80}$$

In addition, the inequality in (79), which is the analogue of (43), assures that (55) is applicable to the solutions $y(t)$ of (80).

Thus, the arguments of Section 7 show that, if δ is sufficiently small in (65), then there is on $t_0 \leqq t < \infty$ a solution $y = y(t)$ of (80) which satisfies the (partial) set of initial conditions (10_1)-(10_2), where $t_0 \geqq T$, and the asymptotic relations (11)-(12) as $t \to \infty$.

In particular, the auxiliary function $w(t)$ of $y = y(t)$ satisfies (60). In view of the fact that $T = T(a, \eta)$ was chosen so large that (68) implies (70) if c^0 satisfies the inequalities in (16) when $c = c_{m+1}$ and c_0 is replaced by $c_0/4 = 1/8(1+\eta)$, it follows from (60) that $w(t) < c_0/4$ for $t_0 \leqq t < \infty$. For, according to the Lemma (and the Remark following its proof) in Section 3, if $w(t) \geqq c_0/4$ for some $t = t^0$ $(\geqq t_0 \geqq T)$, then $w(t) \geqq c_0/8$ for $t \geqq t^0$. This contradicts (60).

Hence, for the solution $y = y(t)$, the functions $M_m(t)$, $L_{ma}(t)$, $N_{m+1}(t)$ satisfy (77). But the defining properties of $F_0(t, y)$ show that this implies $F(t, y(t)) \equiv F_0(t, y(t))$, so that $y = y(t)$ is a solution of (1), as well as of (80). This proves (*bis).

Remark. For applications below, note that the proofs of (*) and (*bis) show that analogues of these assertions remain valid if the "linear part" (2) of (1) is replaced by a suitable system with coefficients depending on t or on (t, y); for example, by a system of the type

$$y^{j1\prime} = \lambda^{j1}(t, y) y^{j1}, \qquad y^{jk\prime} = \lambda^{jk}(t, y) y^{jk} + \delta^{jk}(t, y) y^{j\,k-1},$$

where $\lambda^{jk}(t, y)$, $\delta^{jk}(t, y)$ are complex-valued, continuous functions for large t and small $|y|$. Corresponding to the sets of integers p, q, r in (*), suppose that it is possible to divide the integers $j = 1, \cdots, g$ into three sets, the elements of which will be denoted by p, q, r, with the property that for some pair of constants c and a, where $c > 0$ and $0 < a < 1$,

$$\begin{aligned}
&\operatorname{Re} \lambda^{jk} < -c - a\,|\delta^{jk}| && \text{if } j = q,\\
&\operatorname{Re} \lambda^{pk} + a\,|\delta^{pk}| \leqq \operatorname{Re} \lambda^{qk} - a\,|\delta^{qk}| - c && \text{for all } p, q,\\
&\operatorname{Re} \lambda^{rk} - a\,|\delta^{pk}| \geqq \operatorname{Re} \lambda^{qk} + a\,|\delta^{qk}| + c && \text{for all } q, r,
\end{aligned}$$

where $\delta^{j1} \equiv 0$. Let L_m be defined by (6). Then (*) and (*bis) remain valid if (11) and (12) are replaced, respectively, by

$$|y^{jk}(t)|^2 = o(L_m(t)) \text{ if } j \neq q \quad \text{and} \quad \limsup_{t \to \infty} t^{-1} \log |y(t)| \leqq -c.$$

Part II. Asymptotic formulae.

11. The linear case. Let h_* denote the maximum multiplicity of the elementary divisors of J belonging to characteristic numbers $\lambda(j)$ satisfying $\operatorname{Re} \lambda(j) = \mu$; that is,

$$h_* = \max_q h(q). \tag{81}$$

Let h_0 denote a number (not necessarily an integer) satisfying

$$h_0 \geqq h_*. \tag{82}$$

Finally, let $q = q_0$ denote any fixed value of q, let k_0 be an integer on the range $1 \leqq k_0 \leqq h(q_0)$, and let $b = b(q_0, k_0)$ be defined by

$$b = h_0 - h(q_0) + k_0 \qquad \text{(so that } k_0 \leqq b \leqq h_0\text{)}. \tag{83}$$

As a refinement of a theorem of Dunkel [1], it was shown in [6] that if (1) is a linear system, say $F^j(t, y) = \sum_{k=1}^{d} g_{jk}(t) y^k$, subject to

$$\int^{\infty} t^{h_0 - 1} \, | g_{jk}(t) | \, dt < \infty, \tag{84}$$

then (1) has a solution $y(t) = y_{q_0 k_0}(t)$ the components of which have the following asymptotic properties, as $t \to \infty$: when $q = q_0$,

$$y^{qk}(t) = o(e^{\mu t} t^{k-b}) \text{ if } 1 \leqq k < k, \tag{85$_1$}$$

$$y^{qk}(t) = e^{\lambda(q)t} t^{k-k_0}/(k - k_0)! + o(e^{\mu t} t^{k-b}) \text{ if } k_0 \leqq k \leqq h(q); \tag{85$_2$}$$

when $q \neq q_0$,

$$y^{qk}(t) = o(e^{\mu t} t^{k-b}) \text{ if } 1 \leqq k \leqq h(q); \tag{85$_3$}$$

finally, when $j = p$ or $j = r$,

$$y^{jk}(t) = o(e^{\mu t} t^{1-b}) \text{ if } 1 \leqq k \leqq h(j). \tag{85$_4$}$$

Furthermore, corresponding to every solution $y = y(t)$ of the linear cases of (1) satisfying (12), there exists a set of $\sum h(q)$ constants c^{qk}, not all 0, with the property that $y(t)$ has the following behavior as $t \to \infty$:

$$y(t) = \sum_{q_0} \sum_{k_0} c^{q_0 k_0} y_{q_0 k_0}(t), \tag{86}$$

where $y = y_{q_0 k_0}(t)$ is a vector satisfying the asymptotic formulae (85_1)-(85_4).

Note that if $h_0 \geqq 2h_* - 1$ (for example, if $h_0 \geqq 2d - 1$) in (84), then the remainder terms $o(e^{\mu t} t^{k-b})$ in (85_1)-(85_4) are $o(e^{\mu t})$.

It is clear from (86) that certain terms on the right of the asymptotic formula (86) may be redundant. Thus, for a fixed (qk), the error terms $o(e^{\mu t} t^{k-b})$ are $o(e^{\mu t} t^{k-\beta})$, where $\beta = h_0 - j_0$ and j_0 is the maximum of the differences $h(q_0) - k_0$ for which $c^{q_0 k_0} \neq 0$. Consequently, a "principal" term

$e^{\lambda(q)t}t^{k-k_0}$ of (85) contributes to (86) only if $k_0 \leqq \beta$. The sum (86) can be abbreviated to

$$(87) \qquad y(t) = \sum_q \sum_{k=l(q)}^{k(q)} c^{qk} y_{qk}(t),$$

where $k = l(q)$, $k(q)$ are the least and greatest values of k (if any) satisfying

$$(88) \qquad 1 \leqq l(q) \leqq k(q) \leqq \min(h(q), h_0 - j_0) \text{ and } h(q) - l(q) \geqq j_0.$$

For a given q, the number of terms in the last sum is

$$k(q) - l(q) + 1 \leqq \min(j_0, h_0 - h(q)) + 1.$$

12. Statement of the theorems. In what follows, corresponding assertions will be proved for non-linear systems:

(†) *Let the linear system of differential equations* (2) *with* (*complex-valued*) *constant coefficients have the form* (3). *Let* $\mu = \mu^m$ *satisfy* (8) *and let* (81), (82) *hold. Let* j_0 *be a number satisfying* $0 \leqq j_0 < h_*$ *and* $l(q)$, $k(q)$ *the least, greatest integers, if any, satisfying* (88). *Let* c^{qk}, *where* $l(q) \leqq k \leqq k(q)$, *be given numbers, not all* 0. *Let* $F(t, y)$ *be a vector* (*with complex-valued components*) *defined and continuous for* $0 \leqq t < \infty$, $|y| \leqq \text{const.}$ *Finally, let* $\phi(t, \rho)$ *be a continuous* (*scalar*) *function defined for* $0 \leqq t < \infty$, $0 \leqq \rho \leqq \text{const.}$ *such that* F *and* ϕ *satisfy*

$$(89) \qquad |F(t, y)| \leqq \phi(t, |y|),$$

$$(90) \qquad \phi(t, \rho) \text{ is non-decreasing in } \rho \text{ (for fixed } t\text{)},$$

$$(91) \qquad \int^{\infty} e^{-\mu t} t^{h_0 - j_0 - 1} \phi(t, \text{const.}\, e^{\mu t} t^{j_0})\, dt < \infty \text{ for every const.} > 0.$$

Then (1) *has a solution* $y = y(t)$ *for large* t *satisfying the asymptotic formula* (87).

Conditions (89), (90), (91) are surely satisfied if there exists, for large t, a continuous function $f(t)$ satisfying

$$(94) \qquad |F(t, y)| \leqq f(t)|y| \quad \text{and} \quad \int^{\infty} t^{h_0 - 1} f(t)\, dt < \infty.$$

In fact, $\phi(t, \rho)$ can then be chosen to be $f(t)\rho$.

In the conservative case, where $F(t, y) \equiv F(y)$ does not depend on t, conditions (89), (90), (91) are satisfied if there exists a continuous function $\phi(\rho)$ satisfying the following conditions:

$$|F(y)| \leqq \phi(|y|), \tag{95}$$

$$\phi(\rho) \text{ is non-decreasing,} \tag{96}$$

$$\int_{+0} \rho^{-2} |\log \rho|^{h_0-1} \phi(\rho)\, d\rho < \infty. \tag{97}$$

(In order to verify that (91) is a consequence of (97), introduce the new integration variable t in (97) defined by $\rho = \text{const.}\, e^{\mu t} t^{j_0}$ and note that $\rho^{-1} d\rho \sim \mu dt$ and that $\log \rho \sim \mu t$ as $t \to \infty$ (and/or $\rho \to 0$).)

Assertion (†) has the following partial converse:

(††) *Let* $\mu = \mu^m$ (< 0), h_* *and* h_0 *be defined as in* (†). *Let* $F(t,y)$ *be defined and continuous for* $0 \leqq t < \infty$, $|y| \leqq \text{const.}$ *Let there exist a continuous function* $f(t)$ *for large* t *satisfying* (94) *or let there exist a continuous function* $\phi(t,\rho)$ *for large* t *and small* $\rho \geqq 0$ *satisfying* (89), (90), (91) *and such that*

$$\phi(t,\rho) \text{ is non-increasing in } t \text{ (for fixed } \rho). \tag{90 bis}$$

Let $y = y(t)$ *be a solution of* (1) *for large* t *satisfying* (11) *and* (12). *Then* $y = y(t)$ *satisfies the asymptotic formula* (86) (*that is, the asymptotic formula* (87)-(88) *for some* j_0, *where* $0 \leqq j_0 < h_*$, *and some constants* c^{qk}, *not all* 0).

Under conditions (90) and (90 bis), the assumption (91) is implied by

$$\int_{+0} \rho^{-2} |\log \rho|^{h_0-1} \phi(\nu |\log \rho|, \rho)\, d\rho < \infty \quad \text{for some } \nu < 1/|\mu|. \tag{91 bis}$$

The proof of (††) will consist essentially of two steps: first, of showing that the conditions on $F(t,y)$ and on $y(t)$ imply

$$\int^{\infty} t^{h_0-1} \{|F(t,y(t))|/|y(t)|\}\, dt < \infty;$$

second, of showing that (11), (12) and the last relation imply the asymptotic formula (86) for $y(t)$.

The conditions of (††) on F, that is, either the condition (94) or the set of conditions (89), (90), (90 bis), (91), are sufficient for the validity of both (†) and (††). In particular, in the conservative case, where $F(t,y) \equiv F(y)$, the conditions (95)-(97) are sufficient for the truth of (†) and (††). For variants of these assertions, see Section 18 below.

It will be clear from the proofs that, under condition (94), (†) and

(††) remain valid if $\mu < 0$ is relaxed to $\mu \leqq 0$ (and $F(t, y)$ is defined for all $t \geqq 0$ and all y). The simplest case of these two theorems, where (94) is assumed, J is the zero matrix and $h_0 = 1$, is contained in the theorem of [17]. The latter allows a relaxation of (94), with $h_0 = 1$, to

$$|F(t,y)| \leqq f(t)\phi(|y|), \qquad \int^{\infty} f(t)\,dt < \infty, \qquad \int^{\infty} d\rho/\phi(\rho) = \infty$$

(where $f(t), \phi(\rho) > 0$ are only assumed to be continuous); for a further refinement of this case, see [9], p. 51. It is easy to see, however, that the conditions in the last formula line cannot replace (94) in (††), say, even if $J = \mu I$, where $\mu \neq 0$, I is the d by d unit matrix and $h_0 = 1$.

The binary conservative cases ($d = 2$) of (†), (††), when the real parts of the characteristic roots λ_1, λ_2 of J are equal (and negative), the elementary divisors of J are simple ($h_* = 1$) and $h_0 = 1$, are given in [18]; cf. Section 18 below for a slight refinement. The analogue of the case $\lambda_1 = \lambda_2$ (< 0) and a double elementary divisor is treated in [8], pp. 526-528 (cf. [7]), in the form

$$dy/dx = (y + x + F^2(x, y))/(x + F^1(x, y))$$

under assumptions less severe than (95)-(97), with h_0 $(= h_*) = 2$. Perron [10] had dealt with the general equation

$$(ax + by + F^1(x, y))\,dy = (cx + dy + F^2(x, y))\,dx$$

under the assumption that F^1, F^2 satisfy Lipschitz conditions with Lipschitz constants of the order ρ^ϵ, for some $\epsilon > 0$, on the circle $x^2 + y^2 \leqq \rho^2$; cf. [16].

13. The case $h_0 = h_* = 1$ of (†): Preliminaries. The proof of (†) will be given first for the case $h_* = 1$ (so that the elementary divisors belonging to the eigenvalues λ_k for which $\mathrm{Re}\,\lambda_k = \mu (= \mu^m)$ are simple). In this case, the double sum (22) is a single sum over q (with $k = 1$), and so the differential inequality (25_m) can be replaced by

$$(98) \qquad |L_{ma}' - 2\mu^m L_{ma}| \leqq d\,|F(t, y)|\,|y|.$$

In the proof, it can be supposed that $m > 1$. It will be clear that the case $m = 1$ is substantially simpler (since $v_1(t) \equiv 1$).

Let $y = y(t) \neq 0$ be a solution of (1) satisfying (20) and (46) for some positive c_0 $(< \frac{1}{2})$ (hence (48)) on some interval. Let K be a positive constant so large that

$$(99) \qquad K^{-2}e^{2\mu t} < L_{ma} < K^2 e^{2\mu t}$$

holds on this t-interval. Then, by (89) and (48),

$$|F(t,y(t)| \leqq \phi(t,|y(t)|) \leqq \phi(t,ce^{\mu t}), \text{ where } c^2 = 2a^{-2d}c_0^{-1}K^2. \tag{100}$$

By (48) and the first inequality in (99),

$$|y|/L_{ma} \leqq c/L_{ma}^{\frac{1}{2}} \leqq cKe^{-\mu t}, \tag{101}$$

and so (98), (100), (101) imply

$$|L_{ma}' - 2\mu^m L_{ma}| \leqq cdKe^{-\mu t}\phi(t,ce^{\mu t})L_{ma}. \tag{102}$$

If the interval $[t,t_0]$ is contained in the interval under consideration, a quadrature gives

$$|\log(e^{-2\mu t}L_{ma}(t)/e^{-2\mu t_0}L_{ma}(t_0))| \leqq cdKI(c,t), \tag{103}$$

where

$$I(c,t) = \int_t^\infty e^{-\mu u}\phi(u,ce^{\mu u})\,du. \tag{104}$$

If a, c_0 and K (hence c) are fixed, it follows from (91), where $h_0 - 1 = j_0 = 0$, that there exists a $T = T(a,c_0,K)$ such that (99) implies

$$\tfrac{1}{2} \leqq e^{-2\mu t}L_{ma}(t)/e^{-2\mu t_0}L_{ma}(t_0) \leqq 2, \text{ if } T \leqq t \leqq t_0.$$

This argument shows that if (20), (46) and (99) hold on an interval $t^0 < t \leqq t_0$, where $t^0 \geqq T$, and if

$$2K^{-2}e^{2\mu t_0} < L_{ma}(t_0) < \tfrac{1}{2}K^2e^{2\mu t_0}, \tag{105}$$

then (99) holds at $t = t^0$.

If, in addition, T is sufficiently large and $w(t_0) = 0$, then (46) holds at $t = t^0$. For, by (48), (99), (100) and $|y|^2 \geqq L_{ma}$,

$$|F(t,y(t))| \leqq |y(t)|Ke^{-\mu t}\phi(t,ce^{\mu t}).$$

Since $f(t) = f_K(t) = Ke^{-\mu t}\phi(t,ce^{\mu t})$ satisfies (91), hence (70), the statement concerning (46) at $t = t^0$ is a consequence of the Riccati differential inequality (37) and of the Remark following the Lemma of Section 3.

Consequently, if $w(t) = 0$ and (20), (105) hold at $t = t_0$, then, when t decreases from t_0, either the solution $y(t)$ exists on $T \leqq t \leqq t_0$ and satisfies (20), (46) and (99) on this interval or a point $t = t^0$ $(> T)$ is reached with the properties that (46) and (99) hold on $t^0 \leqq t \leqq t_0$, while (20) holds on $t^0 < t \leqq t_0$ but not at $t = t^0$.

14. The case $h_0 = h_* = 1$ of (†): Completion of the proof. A repetition of the arguments of Sections 7, 8 and 10 shows that *if $T = T(a, c_0, K)$ is sufficiently large, if $t_0 > T$, and if the numbers $y^{qk}(t_0) = y_0^{qk}$ satisfy* (105), *then there exist at least one set of $\Sigma h(p)$ numbers y_0^{pk} and a solution $y = y(t)$ of* (1) *on $T \leqq t \leqq t_0$ satisfying*

$$(106_1) \qquad y^{pk}(t_0) = y_0^{pk}$$

and

$$(106_2) \qquad y^{qk}(t_0) = y_0^{qk}, \qquad (106_3) \qquad y^{rk}(t_0) = 0,$$

finally (20), (46) *and* (99) *on $T \leqq t \leqq t_0$.*

Note that (106_3) implies $w(t_0) = 0$. The fact that the last italicized statement concerns $t \leqq t_0$, while that in Section 7 concerns $t \geqq t_0$, means that the rôles of p, r, of M_m, N_{m+1} and of v, w are interchanged. The "cone" $v \geqq c_0$, that is, $N_m \geqq c_0(N_m + M_m)$ or

$$c_0 \sum_q \sum_k a^{2k} | y^{pk} |^2 \leqq (1 - c_0) (\sum_q \sum_k a^{2k} | y^{qk} |^2 + \sum_r \sum_k a^{2k} | y^{rk} |^2),$$

plays the rôle of (58) in the proof of the desired statement.

Let $y = y(t)$ be a solution of (1) on $T \leqq t \leqq t_0$ supplied by the last italicized statement. Since $h_* = 1$, the index k is 1 in any pair of indices (qk) and the corresponding differential equation for y^{q1} is

$$y^{q1\prime} = \lambda(q) y^{q1} + F^{q1}(t, y).$$

This gives $| (y^{q1} e^{-\lambda(q)t})' | \leqq e^{-\mu t} | F(t, y) | \leqq e^{-\mu t} \phi(t, | y(t) |)$, and so, in view of (48) and (99), $| (y^{q1} e^{-\lambda(q)t})' | \leqq e^{-\mu t} \phi(t, c e^{\mu t})$. Hence a quadrature over $[t, t_0]$, where $T \leqq t \leqq t_0$, gives

$$(107) \qquad | y^{q1}(t) e^{-\lambda(q)t} - y_0^{q1} e^{-\lambda(q)t_0} | \leqq I(c, t); \qquad \text{cf. } (104).$$

The assertion of (†), where $h_0 = 1$, is to the effect that if $\sum_q h(q) = \sum_q 1$ numbers c^{q1}, not all 0, are given, then (1) has a solution satisfying, as $t \to \infty$,

$$(108_1) \quad y^{q1}(t) = c^{q1} e^{\lambda(q)t} + o(e^{\mu t}); \qquad (108_2) \quad y^{jk}(t) = o(e^{\mu t}) \text{ if } j \neq q.$$

In terms of the given numbers c^{q1}, put

$$(109) \qquad y_0^{q1} = c^{q1} e^{\lambda(q) t_0}.$$

When the numbers c^{q1} are fixed, it is possible to choose a fixed, sufficiently large, constant K so that (106_2) satisfies (105).

If $t_0 = n$ is a sufficiently large integer, then, according to the last section, (1) has a solution $y = y_n(t)$ on $T \leqq t \leqq n$, where $T = T(a, c_0, K)$, satisfying (106_2), (106_3) at $t = t_0 (= n)$, and (20), (46) and (99) on $T \leqq t \leqq n$.

The inequality (107), where $T \leqq t \leqq t_0\ (= n)$, is applicable to the solution $y = y_n(t)$. Hence, $|\, y_n^{q1}(t) e^{-\lambda(q)t} - c^{q1} \,| \leqq I(c, t)$. The numbers $|\, y_n(T) \,|$ form therefore a bounded sequence as $n \to \infty$. Hence it is possible to select a subsequence (which can then be supposed to be the full sequence) from $y_1(t), y_2(t), \cdots$ so that

$$y(t) = \lim_{n \to \infty} y_n(t) \tag{110}$$

holds uniformly on every bounded t-interval $T \leqq t \leqq \text{const}$. The limit $y = y(t)$ is a solution of (1) on $T \leqq t < \infty$.

Consequently, for $T \leqq t < \infty$,

$$|\, y^{q1}(t) e^{-\lambda(q)t} - c^{q1} \,| \leqq I(c, t). \tag{111}$$

The definition (104) of $I(c, t)$ and the assumption (91) of (†) show that

$$I(c, t) = o(1) \text{ as } t \to \infty, \tag{112}$$

if c is fixed. The relations (108_1) follow from (111) and (112).

The auxiliary functions $v(t)$, $w(t)$ belonging to $y(t)$ satisfy $v(t) \geqq c_0$, $w(t) \leqq \frac{1}{2} c_0$ for $T \leqq t < \infty$, since the corresponding functions of $y_n(t)$ satisfy (20), (46) for $T \leqq t \leqq t_0\ (= n)$. Hence (11) and (12) hold; cf. Section 5. The relations (88_2) follow from (11), (12) and (108_1). This completes the proof of (†) for the case $h_0\ (= h_*) = 1$.

Remark. For reference below, note that, according to the proof just completed, if $F(t, y)$ is defined for $t \geqq T$, $|\, y \,| \leqq 1 < K(t)$, where $K(T) \to \infty$ as $T \to \infty$, then (†) remains true in the case $h_0 = 1$ when $\mu = 0$ and (89)-(91) are assumed.

15. The case $h_0 > 1$: Change of variables. The proof of (†) in the cases $h_0 > 1$ will depend essentially on the case $h_0 = 1$ and a change of variables similar to that used in the proof of (**) in [6]. Let the independent variable t be replaced by $s = \log t$ (for large t and/or large s),

$$t = e^s \qquad (ds = dt/t). \tag{113}$$

Let the dependent variables be subject to a linear transformation, depending on $t = t(s)$,

$$y = Qz, \text{ where } Q = D_0 D_1 C D_2, \tag{114}$$

and D_0, D_1, C, D_2 are d by d matrices to be given below.

The transformation $y = D_0 u$ is given by

$$y^{qk} = e^{\lambda(q)t} u^{qk} \text{ and } y^{jk} = e^{\mu t} u^{jk} \text{ if } j \neq q. \tag{115}$$

The triangular matrix D_1 depends on the fixed number β, where

$$\beta = h_0 - j_0, \tag{116}$$

and is the matrix of the linear transformation given by

$$u^{qk} = t^{k-\beta} v^{qk} \text{ and } u^{jk} = t^{1-\beta} v^{jk} \text{ if } j \neq q. \tag{117}$$

The matrix C is constant and is the matrix of the linear transformation $v = Cw$, where

$$v^{qk} = \sum_{n=1}^{k} w^{qn}/(k-n)\,! \text{ and } v^{jk} = w^{jk} \text{ if } j \neq q. \tag{118}$$

The change of independent variables (113) and the successive changes (115), (117), (118) of dependent variables transform the system (4) into

$$\dot{w}^{qk} = (\beta - k) w^{qk} + E^{qk}(s, w) \tag{119}$$

and, if $j \neq q$,

$$\begin{aligned} \dot{w}^{j1} &= ((\lambda(j) - \mu)t + \beta - 1) w^{j1} + E^{j1}(s, w), \\ \dot{w}^{jk} &= ((\lambda(j) - \mu)t + \beta - 1) w^{jk} + t w^{j\,k-1} + E^{jk}(s, w), \qquad k > 1, \end{aligned} \tag{120}$$

where a dot denotes differentiation with respect to s and $E(s, w)$ is the vector $t(D_0 D_1 C)^{-1} F(t, D_0 D_1 C w)$, where $t = e^s$. Cf. [6], Sections 12 and 13.

The choice of the transformation $w = D_2 z$ depends on the given j_0 determining $l(q)$ and $k(q)$. In terms of β and the numbers $l(q)$, $k(q)$, define the linear transformation $w = D_2 z$ by

$$\begin{aligned} w^{qk} &= t^{\beta-k} z^{qk} && \text{if } \ l(q) \leqq k \leqq k(q), \\ w^{qk} &= z^{qk} && \text{if } \ k < l(q) \text{ or } k > k(q), \\ w^{jk} &= z^{jk} && \text{if } \ j \neq q. \end{aligned} \tag{121}$$

Since $t = e^s$, (101) transforms the system (119)-(120) into

$$\begin{aligned} \dot{z}^{qk} &= 0 \cdot z^{qk} + B^{qk}(s, z) \quad \text{if } \ l(q) \leqq k \leqq k(q), \\ \dot{z}^{qk} &= (\beta - k) z^{qk} + B^{qk}(s, z) \quad \text{if } \ k < l(q) \text{ or } k > k(q)\,; \end{aligned} \tag{122}$$

and, if $j \neq q$,

$$\begin{aligned} \dot{z}^{j1} &= ((\lambda(j) - \mu)t + \beta - 1) z^{j1} + B^{j1}(s, z), \\ \dot{z}^{jk} &= ((\lambda(j) - \mu)t + \beta - 1) z^{jk} + t z^{j\,k-1} + B^{jk}(s, z) \qquad k > 1, \end{aligned} \tag{123}$$

where the vector $B(s,z)$ is

$$B(s,z)=tQ^{-1}F(t,Qz), \qquad t=e^{s} \tag{124}$$

and Q is the matrix of the transformation $y=Qz$; cf. (114).

By (115), (117), (118) and (121), the transformation $y=Qz$ is given by

$$\begin{aligned} y^{qk}&=e^{\lambda(q)t}\{t^{k-\beta}\Sigma_{\mathrm{I}}\, z^{qi}/(k-i)\,!+\Sigma_{\mathrm{II}}\, t^{k-i}z^{qi}/(k-i)\,!\},\\ y^{jk}&=e^{\mu t}t^{1-\beta}z^{jk} \quad \text{if} \quad j\neq q, \end{aligned} \tag{125}$$

where Σ_{I} indicates the sum over the indices i satisfying $i\leqq k$ and either $i<l(q)$ or $i>k(q)$, and Σ_{II} is the sum over the indices i satisfying $i\leqq k$ and $l(q)\leqq i\leqq k(q)$. The transformation $z=Q^{-1}y$ is given by

$$\begin{aligned} z^{qk}&=e^{-\lambda(q)t}\sum_{i=1}^{k}(-1)^{i}t^{k-i}y^{qi}/(k-i)\,! \text{ if } l(q)\leqq k\leqq k(q),\\ z^{qk}&=e^{-\lambda(q)t}\sum_{i=1}^{k}(-1)^{i}t^{\beta-i}y^{qi}/(k-i)\,! \text{ if } k<l(q) \text{ or } k>k(q),\\ z^{qk}&=e^{-\mu t}t^{\beta-1}y^{jk} \text{ if } j\neq q. \end{aligned} \tag{126}$$

If $|Q|$ denotes the norm of Q (that is, $\max|Qz|$ for $|z|=1$), then (125) shows that $|Q|$ is $O(e^{\mu t})$ times t^{γ}, where $\gamma=\max(h_{*}-\beta, h(q)-l(q))$ and the max refers to the set of q's for which $c^{qk}\neq 0$ for some k. Since $h_{*}-\beta=j_0+h_{*}-h_0\leqq j_0$, it follows from the last part of (88) that

$$|Q|=O(e^{\mu t}t^{j_0}) \quad \text{as} \quad t\rightarrow\infty. \tag{127}$$

By (126), $|Q^{-1}|$ is $O(e^{-\mu t})$ times t^{γ}, where $\gamma=\max(\beta-1, k(q)-1)$. The first part of (88) shows that $k(q)\leqq h_0-j_0\,(=\beta)$. Hence

$$|Q^{-1}|=O(e^{-\mu t}t^{h_0-j_0-1}) \quad \text{as} \quad t\rightarrow\infty. \tag{128}$$

In view of (89) and (124),

$$|B(s,z)|\leqq t\,|Q^{-1}|\,\phi(t,|Qz|). \tag{129}$$

16. Completion of the proof of (†). The relations (127)-(129) give

$$|B(s,z)|\leqq\psi(s,|z|), \tag{130}$$

where, for a sufficiently large Const.,

$$\psi(s,\rho)=\text{Const.}\,te^{-\mu t}t^{h_0-j_0-1}\phi(t, \text{Const.}\,e^{\mu t}t^{j_0}\rho). \tag{131}$$

The function (131) satisfies the condition that

$$(132)\qquad \psi(s,\rho) \text{ is non-decreasing in } \rho \text{ (for fixed } s)$$

and, by (91) and $ds = dt/t$,

$$(133)\qquad \int^{\infty} \psi(s, \text{const.})\, ds < \infty \text{ for every const.} > 0.$$

Conditions (130), (132), (133) are the analogues of (89), (90), (91) for the case $\mu = 0$ and $h_0 - 1 = j_0 = 0$.

The fact that the coefficient of z^{qk} (in the linear terms of the first equation of (122)) is 0, rather than a negative constant, does not prevent the applicability of the case $h_0 = 1$ of (†), under assumptions (89)-(91), to the system (122)-(123); cf. the Remark at the end of Section 14. Nor does the fact that the coefficients of the linear terms on the right of the equations in (123) are not constants prevent the applicability of this assertion, since $\operatorname{Re}(\lambda(j) - \mu) \neq 0$; cf. the Remark at the end of Section 10.

Hence (130), (132) and (133) imply that, for given numbers c^{qk}, not all 0, where $l(q) \leqq k \leqq k(q)$, there is at least one solution $z = z(s)$ of (122)-(123) satisfying, as $s \to \infty$,

$$(134)\qquad \begin{aligned} &z^{qk}(s) \to c^{qk} && \text{if} \quad l(q) \leqq k \leqq k(q), \\ &z^{qk}(s) \to 0 && \text{if} \quad k < l(q) \text{ or } k > k(q), \\ &z^{jk}(s) \to 0 && \text{if} \quad j \neq q. \end{aligned}$$

In view of (125), the corresponding solution $y = y(t)$ of (1) satisfies (87). This proves (†).

17. Proof of (††). Let $y = y(t) \not\equiv 0$ be a solution of (1) for large t. It is clear that, under the assumption (94) on $F(t,y)$, the (vector) function $y = y(t)$ does not vanish and satisfies

$$(135)\qquad \int^{\infty} t^{h_0-1}\{|F(t,y(t))|/|y(t)|\}\, dt < \infty.$$

It will be verified that the same is true if $y = y(t)$ satisfies (11) and (12), where $\mu < 0$, and $F(t,y)$ satisfies (89), (90), (90 bis) and (91).

That $y(t)$ does not vanish (for large t) follows from (12). In view of (11), the auxiliary functions $v(t)$, $w(t)$ satisfy the respective limit relations (44), (60). In particular, (20) and (46) hold, where $0 < c_0 < \frac{1}{2}$, and so (48) and (49) hold for large t. Let

$$(136)\qquad \rho(t) = L_{ma}^{\frac{1}{2}}(t) \ (> 0),$$

so that, by (11), there exist positive constants c, K such that

$$K\rho(t) \leqq |y(t)| \leqq c\rho(t), \tag{137}$$

and, by (12),

$$\log\rho(t) \sim \mu t \quad \text{as} \quad t \to \infty; \tag{138}$$

in particular,

$$t > \nu\,|\log c\rho(t)| \tag{139}$$

if t is sufficiently large and $\nu < 1/|\mu|$.

In view of (89) and (90)-(90 bis), it follows that

$$|F(t, y(t))|/|y(t)| \leqq \phi(\nu\,|\log c\rho(t)|, c\rho(t))/K\rho(t). \tag{140}$$

In order to prove (133), it will first be shown that

$$|F(t, y(t))|/|y(t)| \to 0 \quad \text{as} \quad t \to \infty. \tag{141}$$

To this end, it is sufficient to show that $\phi(\nu\,|\log\rho|, \rho)/\rho \to 0$ as $\rho \to 0$; actually, even more is true, namely,

$$|\log\rho|^{h_0-1}\phi(\nu\,|\log\rho|, \rho)/\rho \to 0 \quad \text{as} \quad \rho \to 0. \tag{142}$$

In order to see this, note that an integration by parts shows that the integral

$$\int_{\rho} \sigma^{-2}\,|\log\sigma|^{h_0-1}\,\psi(\sigma)\,d\sigma,$$

where $\psi(\rho) = \phi(\nu\,|\log\rho|, \rho)$, is

$$\psi(\rho)\int_{\rho} \sigma^{-2}\,|\log\sigma|^{h_0-1}\,d\sigma + \int_{\rho}\Big(\int_{\sigma} \tau^{-2}\,|\log\tau|^{h_0-1}\,d\tau\Big)\,d\psi(\sigma),$$

Since (90)-(90 bis) implies that $d\psi(\sigma) \geqq 0$, it follows that

$$\int_{\rho} \sigma^{-2}\,|\log\sigma|^{h_0-1}\psi(\sigma)\,d\sigma \geqq \psi(\rho)\int_{\rho} \sigma^{-2}\,|\log\sigma|^{h_0-1}\,d\sigma.$$

The factor of $\psi(\rho)$ on the right is $O(\rho^{-1}\,|\log\rho|^{h_0-1})$ as $\rho \to 0$. Thus (142) follows from (91 bis) (which, in turn, is a consequence of (91)).

It is seen from (141) that

$$d\rho/dt < 0 \text{ for large } t; \tag{143}$$

in fact, for large t,

$$(\mu - 2a)\rho < d\rho/dt < (\mu + 2a)\rho, \tag{144}$$

by (49) and (141). The inequality (143) assures that it is possible to introduce the new integration variable $\rho = \rho(t)$ into

(145) $$\int^{\infty} \rho^{-1}(t) t^{h_0-1} \phi(\nu \,|\log c\rho(t)|, c\rho(t))\, dt.$$

In view of (91 bis), (138) and (144), the integral (145) is finite. Consequently, (135) follows from (140).

In terms of the vectors $F = (F^1, \cdots, F^d)$ and $y(t) = (y^1(t), \cdots, y^d(t))$, define, for large t, a d by d matrix $G(t) = \| g_{jk}(t) \|$ as follows:

(146) $$g_{jk}(t) = F^j(t, y(t)) \bar{y}^k(t) / | y(t) |^2,$$

where the bar denotes complex conjugation. The function (146) is continuous for large t and, by (135), satisfies condition (84). Since $y = y(t)$ is a solution of (1), it satisfies the linear system

(147) $$y' = (J + G(t)) y.$$

Hence, the assertion (86) of (††) follows from the theorem (**) of [6] on the linear systems described in Section 11 (or, equivalently, from (†) above, applied to (147), and from the superposition principle). This proves (††).

18. Variants of (†) and (††). The variant of (†) to be considered implies that in the conservative case, $F(t, y) \equiv F(y)$, of (†) the condition (96) can be relaxed to

(148) $$| F(y) | / | y | \to 0 \quad \text{as} \quad | y | \to 0$$

(whereas (95)-(97) imply the stronger condition

(149) $$(\log | y |)^{h_0-1} | F(y) | / | y | \to 0 \quad \text{as} \quad | y | \to 0;$$

cf. the deduction of (142)). It can be remarked, however, that a disadvantage of the requirement of the existence of a $\phi(\rho)$ satisfying the set of conditions (95), (148) and (97), rather than the set (95), (96) and (97), is that the first set is not invariant under affine transformations of the variables y.

(† bis) *Assertion* (†) *remains valid if the conditions* (90), (91) *on* $\phi(t, \rho)$ *are replaced by the conditions* (7) *on* $F(t, y)$ *and the conditions* (90 bis), (91 bis) *on* $\phi(t, \rho)$.

It will be clear from the proof that (7) can be relaxed to the requirement that

(150) $$| F(t, y) | / | y | \to 0 \quad \text{as} \quad t \to \infty$$

uniformly on every (t, y)-set

$$K^{-1}e^{\mu t}t^{j_0} \leqq |y| \leqq K e^{\mu t} t^{j_0}, \tag{151}$$

where $K > 0$ is an arbitrary constant.

In view of († bis), there arises the question whether or not conditions (90), (91) in (††) can be replaced by conditions (7), (90 bis), (91 bis). This question will be answered here only in certain cases.

(†† bis) *Suppose that either* $h_* = 1$ *or* $\mu < -1$. *Then* (††) *remains valid when condition* (90) *is omitted and condition* (91) *is replaced by* (7) *and* (91 bis).

The number 1 in the inequality $\mu < -1$ is the number occurring on the subdiagonal of J.

19. Proof of († bis). Consider the change of variables (113), (114) in (1), leading to the system (122)-(123). For any solution $z = z(s)$ of (122)-(123), let $y = y(t)$ denote the solution of (1) given by $t = e^s$ and $y = Qz$.

It will first be shown that, *under condition* (7), *when a and $1/K$ are fixed (sufficiently small) positive numbers, there exist positive numbers $T = T(a, K)$ and $\epsilon = \epsilon(a)$ with the property that if $z = z(s)$ is a solution of* (122)-(123) *on some s-interval, where $t = e^s \geqq T$, satisfying*

(152) $|z^{qk} - c^{qk}| < \epsilon$ if $l(q) \leqq k \leqq k(q)$; (153) $|z^{jk}| < \epsilon$ otherwise,

then the corresponding solution $y = y(t)$ of (1) *satisfies*

(154) $K^{-1}e^{\mu t}t^{j_0} < |y| < K e^{\mu t} t^{j_0}$, (155) $\big|\,|y|' - \mu |y|\,\big| \leqq a|y|$.

In (154), j_0 denotes the maximum of the differences $h(q) - k$ for which $c^{qk} \neq 0$; so that it can be supposed that $j_0 = h(q) - l(q)$ for some q. Let $q = q_0$ denote any q satisfying $j_0 = h(q) - l(q)$. Let

$$L^0 = \Sigma^0 |y^{qh}|^2, \tag{156}$$

where Σ^0 denotes a sum over those $q = q_0$. Let K be so large that (152), (153) imply

$$(\tfrac{1}{2}K)^{-1} < \Sigma^0 |z^{qk}|^2 \leqq |z|^2 < \tfrac{1}{2}K. \tag{157}$$

Then (125), where $h_* - \beta \leqq j_0$, and (152) show that

$$|y^{jk}| \leqq \epsilon e^{\mu t} t^{j_0} \text{ if } j \neq q,$$

$$|y^{qk}| \leqq \tfrac{1}{2}K d e^{\mu t} t^{j_0 - 1} \text{ if } q = q_0,\ k < h \text{ or if } q \neq q_0,\ k \leqq h,$$

and that $|y^{qh}(t)|$, when $q = q_0$, differs from $|c^{qh}|e^{\mu t}t^{j_0}/(h-l)!$ by terms which are majorized either by $\epsilon e^{\mu t}t^{j_0}$ or by $Ke^{\mu t}t^{j_0-1}$. Hence, if $\epsilon = \epsilon(a)$ is sufficiently small, $T = T(a, K)$ sufficiently large and $t \geqq T$, then (154) holds and

$$L^0(t) \leqq |y(t)|^2 \leqq (1+a^2)L^0(t). \tag{158}$$

From the differential system (4) and from $(|y^{jk}|^2)' = \bar{y}^{jk}y^{jk\prime} + \bar{y}^{jk\prime}y^{jk}$, it follows that $||y|^{2\prime} - 2\mu L_0(t)|$ is majorized by a constant multiple of the sum of $|y(t)|^2 - L^0(t)$, $|y(t)|(|y(t)|^2 - L^0(t))^{\frac{1}{2}}$ and $F(t, y(t))| \cdot |y(t)|$. In view of (7) and (158), the inequality (155) follows if a is sufficiently small and T $(\geqq t)$ is sufficiently large. This proves the italicized statement concerning (152)-(155).

The proof of († bis) can now be completed along the lines used in Sections 13-14 in the proof of (†) when $h_0 = 1$. To this end, define L_a, $M = M_{ma}$, $N = N_{m+1\,a}$ by the formulae

$$L_a = a^2 \sum_q \sum_{k=l(q)}^{k(q)} |z^{qk}|^2, \tag{159}$$

$$M = \sum_p \sum_k a^{2k} |z^{pk}|^2 + a^2 \sum_q \sum_{k>k(q)} |z^{qk}|^2, \tag{160}$$

$$N = \sum_r \sum_k a^{2k} |z^{pk}|^2 + a^2 \sum_q \sum_{k<l(q)} |z^{qk}|^2. \tag{161}$$

Thus, the analogues of the auxiliary functions $v = v(s)$, $w = w(s)$ are

$$v = (N + L_a)/(M + L_a + N), \qquad w = N/(M + L_a + N). \tag{162}$$

These satisfy the Riccati differential inequality

$$\dot{u} \geqq \tfrac{1}{2}u(1-u) - d^2a^{-2d}B(s, z(s))|/|z(s)|, \qquad (u = v \text{ or } u = w); \tag{163}$$

cf. the derivation of (30), (33).

Consider a solution $z = z(s)$ of (122)-(123) satisfying the (possibly incomplete) set of initial conditions

$$z^{qk}(s_0) = c^{qk} \text{ if } l(q) \leqq k \leqq k(q); \qquad N(s_0) = 0. \tag{164}$$

Suppose further that (152) and that (20) and (46) hold on some s-interval (s, s_0) for some c_0, where $0 < c_0 < \frac{1}{2}$. Then the analogue of (47) gives

$$M + N < (2c_0^{-1} - 1)L_a \tag{165}$$

on this interval. Thus, if $c_0 = c_0(a)$ is sufficiently near $\frac{1}{2}$, the inequalities (152) on $(s, s_0]$ imply (153) there.

A quadrature of the equations (122) shows that

$$(166) \qquad | z^{qk}(s) - c^{qk} | \leqq \int_{s}^{s_0} | B(s, t(s)) | \, ds, \qquad l(q) \leqq k \leqq k(q).$$

In view of (128), (129) and $ds = dt/t$,

$$\int_{s}^{s_0} | B(s, z(s)) | \, ds \leqq \text{Const.} \int_{t}^{t_0} e^{-\mu t} t^{h_0 - j_0 - 1} \phi(t, | y(t) |) \, dt,$$

where $t_0 = e^{s_0}$, $t = e^{s}$.

By virtue of the last italicized statement, (154)-(155) hold on the interval under consideration. If a is so small that $\mu + a < 0$, then, by (155), $| y(t) |$ is decreasing on the interval under consideration, and so it is possible to introduce $\rho = | y(t) |$ as a new integration variable in the last integral. By (155), $dt \leqq | \mu + a |^{-1} | d\rho | / \rho$ and, by (154), $e^{-\mu t} t^{-j_0} < K/\rho$. In addition, $t > \nu | \log \rho |$ for $\nu = 1/| \mu - a |$ and $t \geqq T$ if $T = T(a, K)$ is sufficiently large. Hence

$$(167) \qquad \int_{s}^{s_0} | B(s, z(s)) | \, ds \leqq cI(Ke^{\mu t} t^{j_0}),$$

where $c = c(a, K)$ is a constant and

$$(168) \qquad I(\sigma) = \int_{+0}^{\sigma} \rho^{-2} | \log \rho |^{h_0 - 1} \phi(\nu | \log \rho |, \rho) \, d\rho.$$

It follows from (166), (167) and (168), that if (152), (20) and (46) hold on an interval $s^0 < s \leqq s_0$ (that is, $t^0 < t \leqq t_0$) and if t^0 exceeds a sufficiently large $T = T(a, k)$, then (152) also holds at $s = s^0$ (that is, at $t = t^0$). In view of the case $u = w$ of (163), where $| z | > 1/K$ (cf. (157)) on $t^0 < t \leqq t_0$, and of (167), (168), the Lemma of Section 3 implies that if $T = T(a, c_0, K)$ is suitably chosen and $t^0 \geqq T$, then (46) holds at t^0, since $w(t_0) = 0$, by (164).

These facts and the estimates (166), (167) show that the proof of († bis) can now be completed by the arguments of Section 14 used to complete the proof of the case $h_0 = 1$ of (†).

20. **Proof of (†† bis).** It is clear from the proof of (††) that in order to verify the variant of (††), it is sufficient to show that, under the conditions of these statements, (135) holds or, what is more, that

$$(169) \qquad \int^{\infty} t^{h_0-1}\phi(t, |\, y(t)\,|)\, dt < \infty.$$

To this end, it is enough to show that, for large t,

$$(170) \qquad |\, y(t)\,|' < -\text{const.}\, |\, y(t)\,|, \qquad (\text{const.} > 0).$$

For this means that it is possible to introduce the new integration variable $\rho = |\, y(t)\,|$ in the last integral and that $dt \leqq \text{const.}\, |\, d\rho\,|/\rho$. This fact, when combined with

$$(171) \qquad \log |\, y(t)\,| \sim \mu t \quad \text{as} \quad t \to \infty$$

(cf. (11), (12)), and assumption (91 bis) imply (169).

Condition (11) means that, as $t \to \infty$,

$$(172) \qquad |\, y(t)\,|^2 \sim L(t),$$

where $L = L_m$ is defined in (6). It is clear from the differential equations (4) that $|(|\, y(t)\,|^2)' - 2\mu L(t)|$ is majorized by

$$\text{const.}\, \{(|\, y(t)\,|^2 - L(t)) + |\, F(t,y)\,|\, |\, y\,|\} + 2 \sum_q \sum_{k=2}^{h} |\, y^{qk} y^{q\, k-1}\,|.$$

If $h_* = 1$, the last term is missing. If $h_* \geqq 1$, the last term is not greater than $2L(t)$. Hence (170), for large t, follows from (172) if either $h_* = 1$ or $\mu < -1$. This proves (†† bis).

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] O. Dunkel, "Regular singular points of a system of homogeneous linear differential equations of the first order," *Proceedings of the American Academy of Arts and Sciences*, vol. 38 (1902-03), pp. 341-370.

[2] P. Hartman, "Unrestricted solution fields of almost separable differential equations," *Transactions of the American Mathematical Society*, vol. 63 (1948), pp. 560-580.

[3] ——— and A. Wintner, "On the asymptotic behavior of the solutions of a non

linear differential equation," *American Journal of Mathematics*, vol. 68 (1946), pp. 301-308.

[4] ——— and A. Wintner, "On the behavior of the solutions of real binary differential systems at singular points," *ibid.*, vol. 75 (1953), pp. 117-126.

[5] ——— and A. Wintner, "On curves defined by binary non-conservative differential systems," *ibid.*, vol. 76 (1954), pp. 497-501.

[6] ——— and A. Wintner, "Asymptotic integrations of linear differential equations," *ibid.*, vol. 77 (1955), pp. 45-87.

[7] E. Lonn, "Knoteninvarianz bei Differentialgleichungen," *Jahresbericht der Deutschen Mathematiker-Vereinigung;* vol. 43 (1934), pp. 232-237.

[8] ———, "Über singuläre Punkte gewöhnlicher Differentialgleichungen," *Mathematische Zeitschrift*, vol. 44 (1939), pp. 507-530.

[9] S. Łojasiewicz, "Sur l'allure asymptotique des intégrales du système d'équations différentielles au voisinage de point singulier," *Annales Polonici Mathematici*, vol. 1 (1954), pp. 34-72.

[10] O. Perron, "Ueber die Gestalt der Integralkurven einer Differentialgleichung erster Ordnung in der Umgebung eines singulären Punktes," *Mathematische Zeitschrift*, vol. 15 (1922), pp. 121-146 and vol. 16 (1923), pp. 273-295.

[11] ———, "Über Stabilität und asymptotisches Verhalten der Integrale von Differentialgleichungssystemen," *ibid.*, vol. 29 (1929), pp. 129-160.

[12] ———, "Die Stabilitätsfrage bei Differentialgleichungen," *ibid.*, vol. 32 (1930), pp. 703-728.

[13] Z. Szmydtówna, "Sur l'allure asymptotique des intégrales des équations différentielles ordinaires," *Annales de la Société Polonaise de Mathématique*, vol. 24 (1951), No. 2, pp. 17-34.

[14] T. Ważewski, "Sur un principe topologique de l'examen de l'allure asymptotique des intégrales des équations différentielles ordinaires," *ibid.*, vol. 20 (1947), pp. 279-313.

[15] ———, "Sur les intégrales d'un système d'équations différentielles tangentes aux hyperplans caractéristiques issus du point singulier," *ibid.*, vol. 21 (1948), pp. 277-297.

[16] A. Wintner, "Asymptotic integration constants in the singularity of Briot-Bouquet," *American Journal of Mathematics*, vol. 68 (1946), pp. 293-300.

[17] ———, "Asymptotic integrations constants," *ibid.*, vol. 68 (1946), pp. 553-559.

[18] ———, "Vortices and modes," *ibid.*, vol. 69 (1947), pp. 816-824.

PRIMITIVE ALTERNATIVE RINGS AND SEMI-SIMPLICITY.*

By ERWIN KLEINFELD.

---

1. Introduction. Jacobson has characterized primitive, associative rings as dense rings of linear transformations on vector spaces over division rings and developed a structure theory for semi-simple associative rings without chain conditions [3], which generalizes the well known Wedderburn theorems. It seems natural enough to ask whether this theory extends to alternative rings. This question was first proposed by Smiley ([8], question 5). Kaplansky has shown [4] that $\mathfrak{p}$-regular, primitive, alternative rings are either associative or Cayley-Dickson algebras. He also conjectured that the assumption of $\mathfrak{p}$-regularity would prove to be superfluous. The present note shows that indeed this is the case.

The method of attack is to prove a primitive, alternative ring either associative or simple. This in turn is facilitated by the ease with which one can construct two-sided ideals within right ideals. As a by-product we obtain an elementary proof that an alternative, non-associative ring without proper two-sided ideals has no proper one-sided ideals. This accentuates how different the associative theory is from the alternative, for while it is an elementary exercise to show that an associative ring without proper right ideals is either trivial or a division ring, the last mentioned result does not seem to make the classification of simple, alternative rings any easier. Indeed it is still an open question whether alternative, non-associative nil rings without proper ideals exist.

The notion of primitive ring lends itself to a structure theory for semi-simple rings. Every alternative ring modulo its radical turns out to be a subdirect sum of primitive rings.

2. Definitions and identities. In order to facilitate the reading of the proofs we mention some of the essential definitions and identities. For background material as well as development of the identities we refer the reader to [2].

We define $(a,b,c) = ab\cdot c - a\cdot bc$, and $(a,b) = ab - ba$, for arbitrary elements, a, b, c of a ring R. If R is assumed to be alternative then (a,b,c) becomes a skew-symmetric function in its three variables.

The function $f(w,x,y,z)$ is defined by the equation

$$(wx, y, z) = f(w,x,y,z) + x(w,y,z) + (x,y,z)w.$$

---

* Received May 25, 1955.

$f(w,x,y,z)$ turns out to be skew-symmetric in its four variables. Furthermore it satisfies the equation $f(w,x,y,z)=((w,x),y,z)+(w,x,(y,z))$.

The nucleus N of R is defined as the set of all n in R with the property $(n,R,R)=0$. The center C of R is defined as the set of all c in N with the property $(c,R)=0$. The connection between N and C is explained by Lemma 2.3 of [5].

An alternative ring is defined to be simple in case it has no proper two-sided ideals and is not a nil ring.

An alternative ring is defined to be primitive in case it contains a regular, maximal right ideal, which contains no two-sided ideal of the ring other than the zero ideal.

The radical of an alternative ring is defined as the intersection of its regular, maximal right ideals.

3. Construction of ideals. If A is any right ideal of an alternative ring R then we define an element s to be special relative to A in case (i) s is an element of A and (ii) $Rs \subset A$. The following Lemma justifies this definition.

LEMMA 1. *For any right ideal A of an alternative ring R, the special elements form an ideal of R.*

Proof. Let s, s' be any two special elements. It follows directly from the definition that $s-s'$ must then be special. If x and y are arbitrary elements of R, consider first xs. By definition xs is in A, while $y\cdot xs=-(y,x,s)+yx\cdot s$. Since A is a right ideal of R and s is a member of A, it follows that $(y,x,s)=(s,y,x)=sy\cdot x-s\cdot yx$ is also in A. We have shown that $y\cdot xs$ is an element of A and consequently xs is special. So far it has been established that the special elements form a left ideal of R. But also $y\cdot sx=-(y,s,x)+ys\cdot x$. Since ys is in A, so is $ys\cdot x$ and therefore $y\cdot sx$. But then the special elements form a two-sided ideal, as was to be shown.

In the next Lemma we develop an important property of the ideal of special elements.

LEMMA 2. *For any right ideal A of an alternative ring R all elements of the form (A,A,R) are special.*

Proof. As before $(A,A,R)\subset A$. Let us consider elements of the form $R(A,A,R)$. Select any a, a' in A and r, s in R. Then

$$s(r,a,a')=(rs,a,a')-f(r,s,a,a')-(s,a,a')r.$$

Clearly (rs, a, a') is in A, as well as $(s, a, a')r$. Since

$$f(r, s, a, a') = ((r, s), a, a') + (r, s, (a, a')),$$

we obtain also that $f(r, s, a, a')$ is a member of A. Consequently $R(A, A, R) \subset A$, which makes all elements of the form (A, A, R) special.

At this point we obtain the following result:

THEOREM 1. *If R is an alternative, non-associative ring which has no proper two-sided ideals, then it has no proper left or right ideals.*

Proof. Let us suppose that A is a proper right ideal of R. Then by virtue of Lemma 1 the special elements of A consist of just the zero element. Then we may use Lemma 2 to get $(A, A, R) = 0$. From Theorem 3.2 of [5] it follows that the nucleus N and the center C of R coincide. This in turn is used to prove Lemma 3.3 of [5], which asserts that if $(x, y, R) = 0$ then $(x, y) = 0$. From this we conclude that $(A, A) = 0$. In other words A must be an associative, commutative subring of R. For any elements r, s in R and a, a' in A,

$$f(r, s, a, a') = (rs, a, a') - s(r, a, a') - (s, a, a')r = 0,$$

thereby making $f(A, A, R, R) = 0$. As a consequence of this

$$0 = (ar, s, a') = f(a, r, s, a') + r(a, s, a') + (r, s, a')a,$$

so that $(r, s, a')a = 0$. This implies that $(R, R, A)A = 0$. Since $(R, R, A) \subset A$, while A is commutative, also $A(R, R, A) = 0$. But then

$$(aa', r, s) = f(a, a', r, s) + a'(a, r, s) + (a', r, s)a = 0.$$

Consequently $(A^2, R, R) = 0$ and A^2 is in the center of R. In that case $RA^2 = A^2R \subset A$, and hence all elements of the form A^2 are special. This of course implies again $A^2 = 0$. Now $A + RA$ is a two-sided ideal of R. Since it contains $A \neq 0$, it must be all of R. But then $RA = (A + RA)A = 0$. Therefore A is a two-sided ideal of R. But this contradicts the original assumption that R has no proper two-sided ideal and the proof is completed.

This result is new only to the extent that it applies to nil rings. For we have shown elsewhere [6] that alternative, non-associative, simple rings are Cayley-Dickson algebras. The present proof however is quite elementary. In fact we may use it to establish the well known result that Cayley-Dickson algebras have no proper left or right ideals. The question of whether alternative, non-associative, nil rings without proper ideals exist is an open one. For aesthetic reasons one would like to be able to answer this in the negative.

4. Primitive rings. Throughout this section R will denote an alternative, non-associative, primitive ring and A a regular, maximal right ideal of R, which contains no two-sided ideal of R other than the zero ideal. Again it is clear that the special elements of A are all zero, so that $(A, A, R) = 0$ and $(R, R, A)A = 0$, using the same arguments as before. Next we prove

LEMMA 3. *If an element t of R has the property $Rt = 0$, then $t = 0$.*

Proof. The regularity of A assures the existence of an element g in R with the property that for every x in R $gx - x$ is in A. Then in particular $gt - t$ is in A. But $gt = 0$, so that t itself must be in A. In that case t is special relative to A and $t = 0$.

The next step is to show that R is prime.

LEMMA 4. *If B and B' are ideals of R such that $BB' = 0$, then either $B = 0$ or $B' = 0$.*

Proof. Suppose that $B \neq 0$. Then B cannot be contained in A. Since A is a maximal, right ideal and $A + B$ is a right ideal properly containing A, we must have $A + B = R$. Then $RB' = (A + B)B' = AB' + BB' = AB' \subset A$. But clearly RB' is an ideal of R, so that $RB' = 0$. Lemma 3 now tells us that $B' = 0$, which is what we wanted to show.

Denote by N the nucleus of R and by C the center of R. Then with the aid of Lemma 2.3 of [5] we prove the following results.

LEMMA 5. (i) *$N = C$.* (ii) *if a, b are elements of R such that $(a, b, R) = 0$, then $(a, b) = 0$.*

Proof. Suppose S is any non-zero subset of R and suppose the two-sided annihilators of S form an ideal T of R. Then T must be the zero ideal. For since T is an ideal, the two-sided annihilators of it also form an ideal, which moreover contains S. Then by virtue of Lemma 4, T must be zero.

If we assume $N \neq C$ then $(N, R) \neq 0$. Since $(N, R) \subset N$, the annihilators of (N, R) form an ideal, which must then be the zero ideal. From this point on we appeal to [5]. Let x be an arbitrary element of R. Then every element of the form $((R, R, R), R, R)$ or of the form $(x, R, R)(x, R, R)$ is in this ideal. The latter leads us to conclude that if u is any element of the form (R, R, R), then u is an element of N and $u^2 = 0$. The annihilators of (u, R) form an ideal, which contains u since $(x, R, R)(x, R, R) = 0$. If $u \neq 0$ then $(u, R) = 0$, so in either case u is an element of C. Note that Ru is an ideal of R and that $(Ru)^2 = 0$. Using Lemma 4 we obtain $Ru = 0$, so that with the help of Lemma 3 $u = 0$. Since the argument holds for any associator u, R must be associative, which is contrary to assumption. The contradiction arose from the assumption that $N \neq C$, hence (i) is established.

Let c, c' be arbitrary elements of C. Then Rc and Rc' are ideals of R. If in addition $cc' = 0$, then $(Rc)(Rc') = 0$ and as before either $Rc = 0$ or $Rc' = 0$. Consequently either $c = 0$ or $c' = 0$. In other words C has no divisors of zero. Assume now that $(a, b, R) = 0$. Then as before it follows that $f(a, b, R, R) = 0$, and then $((a, b), R, R) = 0$, which means that (a, b) belongs to the set $N = C$. But $(a, ab, R) = 0$, so that $(a, ab) = a(a, b)$ also belongs to C. Then $0 = (a(a, b), b) = (a, b)^2$, and from the previous remark we may conclude that $(a, b) = 0$. This concludes the proof of the lemma.

Now for the main result:

THEOREM 2. *Every primitive, alternative, non-associative ring is a Cayley-Dickson algebra.*

Proof. Using $(A, A, R) = 0$, $(A, R, R)A = 0$ and Lemma 5 we obtain $(A, A) = 0$, $f(A, A, R, R) = 0$, $(A^2, R, R) = 0$ and $A^2 \subset C$. Suppose c belongs to the intersection of A and C. Then $Rc = cR \subset A$. Since Rc is an ideal of R, we get $Rc = 0$, which implies $c = 0$. Since all elements of the form A^2 belong to this intersection, $A^2 = 0$. Again either A is an ideal of R, in which case $A = 0$, or $A + RA$, which is an ideal of R, contains A properly and so $A + RA = R$. But if the latter is true then $RA = (A + RA)A = 0$, so in any case $A = 0$. As A was chosen to be a maximal right ideal of R this simply means that R has no proper right ideals. Also g now becomes a left unit of R and hence R cannot be nil. Then R is a simple ring and therefore a Cayley-Dickson algebra.

5. Semi-simple rings. Having defined the radical Q of an alternative ring R as the intersection of its regular, maximal right ideals, we shall call R semi-simple in case $Q = 0$. An ideal of R will be said to be primitive in case R modulo this ideal is a primitive ring. It is clear that the intersection of the primitive ideals is contained in Q and we proceed to prove equality using our knowledge of primitive rings. It will suffice to show that in a primitive ring the intersection of all regular, maximal right ideals is zero. We have already seen that Cayley-Dickson algebras contain no proper right ideals, while it follows from Jacobson's density theorem for associative, primitive rings that the intersection of all regular, maximal right ideals is zero. For take a fixed representation of R as a dense ring of linear transformations. Then the annihilators of any non-zero vector α form a regular, maximal right ideal (an element sending a into a acting as left unit) and the intersection of all of these is clearly zero. We have proved

THEOREM 3. *The radical Q of any alternative ring R is a two-sided ideal having the property that R/Q is a subdirect sum of Cayley-Dickson algebras and primitive, associative rings.*

The radical Q has also been characterized by Brown [1] as the set of all elements a in R such that every element b in the ideal generated by a has the property that $(1-b)R$ generates R as a right ideal. Kaplansky has shown [4] that the radical defined by Smiley [7] is always contained in Q and coincides with Q in the case of a Zorn ring. Whether it coincides in an arbitrary alternative ring R is an open question. No general structure theory for R modulo its Smiley radical seems to be known.

In conclusion we point out some immediate applications of Theorem 3. Various associativity conditions that hold for Cayley-Dickson algebras now carry over to semi-simple alternative rings. For example it follows that the square of every associator of a semi-simple alternative ring lies in its center while the square of every commutator lies in its nucleus. For arbitrary alternative rings it is known only that the fourth power of every commutator lies in its nucleus [6].

OHIO STATE UNIVERSITY.

REFERENCES.

[1] B. Brown, "An extension of the Jacobson radical," *Proceedings of the American Mathematical Society*, vol. 2 (1951), pp. 114-117.

[2] R. H. Bruck and E. Kleinfeld, "The structure of alternative division rings," *Proceedings of the American Mathematical Society*, vol. 2 (1951), pp. 878-890.

[3] N. Jacobson, "The radical and semi-simplicity for arbitrary rings," *American Journal of Mathematics*, vol. 67 (1945), pp. 300-320.

[4] I. Kaplansky, "Semi-simple alternative rings," *Portugaliae Mathematica*, vol. 10 (1951), pp. 37-50.

[5] E. Kleinfeld, "An extension of the theorem on alternative division rings," *Proceedings of the American Mathematical Society*, vol. 3 (1952), pp. 348-351.

[6] ———, "Simple alternative rings," *Annals of Mathematics*, vol. 58 (1953), pp. 544-547.

[7] M. F. Smiley, "The radical of an alternative ring," *Annals of Mathematics*, vol. 49 (1948), pp. 702-709.

[8] ———, "Some questions concerning alternative rings," *Bulletin of the American Mathematical Society*, vol. 57 (1951), pp. 36-43.

[9] M. Zorn, "Theorie der alternativen Ringe," *Abhandlungen des Mathematischen Seminars der Hamburgischen Universität*, vol. 8 (1930), pp. 123-147.

APPENDIX TO THE PAPER "MEAN PERIODIC FUNCTIONS I".*

By LEON EHRENPREIS**

In our paper, "Mean Periodic Functions I" (see [2]) we considered the problem of whether an entire function f of n complex variables can be approximated by linear combinations of the exponential monomials which occur as the limits of its translates. The following result was proven: Let $\mathcal{H}$ denote the space of entire functions of n complex variables with its usual topology (see [3]), and let V be any variety in $\mathcal{H}$, i. e., V is a closed linear subspace of $\mathcal{H}$ with $V \neq \{0\}$ which is closed with respect to translation. Denote by $\mathcal{H}'$ the dual of $\mathcal{H}$ with the topology of uniform convergence on the bounded sets of $\mathcal{H}$; $\mathcal{H}'$ is a commutative topological ring under convolution. Call V' the set of all $S \varepsilon \mathcal{H}'$ such that $S \cdot f = 0$ for all $f \varepsilon V$. It is readily verified that V' is a closed ideal in $\mathcal{H}'$ which we call the annihilator ideal of V. Then we showed that if V' is a principal ideal (that is, V' has one generator as a closed ideal) then every $f \varepsilon V$ is the limit of the linear combinations of the exponential *monomials* in V. In the present paper we shall show that such is not the case if V' is not principal. (See [2], Problem 2, p. 306).

We showed in [2] that the question of whether or not every variety V is the closure of the linear combinations of the exponential monomials in V is equivalent to the following problem on ideals in rings of analytic functions: Denote by $\boldsymbol{H}'$ the Fourier transform of $\mathcal{H}'$, so $\boldsymbol{H}'$ is the ring of all entire functions of exponential type of n complex variables $Z_1, Z_2, \cdots, Z_n$; the topology of $\boldsymbol{H}'$ is described in [2]. Let I be a closed ideal in $\boldsymbol{H}'$; is I determined by its cospectrum (see [2])? That is, suppose $F \varepsilon \boldsymbol{H}'$ has the property that, for every point z in complex Euclidean n space and any non-negative integers $r_1, r_2, \cdots, r_n$, the conditions

$$(1) \qquad [(\partial^{r_1'+r_2'+\cdots+r_n'}/\partial Z_1^{r_1'} \partial Z_2^{r_2'} \cdots \partial Z_n^{r_n'})H](z) = 0$$

for all $H \varepsilon I$ whenever $0 \leqq r_1' \leqq r_1, 0 \leqq r_2' \leqq r_2, \cdots, 0 \leqq r_n' \leqq r_n$ imply

$$(2) \qquad [(\partial^{r_1+r_2+\cdots+r_n}/\partial Z_1^{r_1} \partial Z_2^{r_2} \cdots \partial Z_n^{r_n})F](z) = 0;$$

is $F \varepsilon I$? We shall produce a closed ideal I in $\boldsymbol{H}'$ and a function $F \varepsilon \boldsymbol{H}'$ such that conditions (1) imply (2), but $F \notin I$.

---

* Received July 22, 1955.

** Work supported by National Science Foundation Grant NSF5-G1010.

(At this point we should like to note some error in the terminology used in [2]. We have often used the term "exponential polynomial in V" for "linear combination of the exponential monomials of V." The spectrum of a variety V as defined in [2] corresponds to the exponential monomials in V. Thus, problem 2 on p. 306 should be formulated as follows: "Is every variety determined by its exponential monomials, that is, if W is a variety and $f \varepsilon W$, is f the limit of the linear combinations of the exponential monomials of W?" This is equivalent to the statement about cospectra of ideals made there. Thus, our proof of Theorem 9 on p. 321 really proves that V is determined by its exponential *monomials*—a much stronger statement than that made in Theorem 9. The reader will have no difficulty in correcting the other cases of this error in notation.)

Indeed, let $n = z$, and let I be the closed ideal in $\boldsymbol{H}'$ generated by the functions $H_1(z_1, z_2) = z_1 + z_2$, $H_2(z_1, z_2) = z_1 z_2$. Since the topology of H' is stronger than pointwise convergence, every $H \varepsilon I$ must vanish at the origin. On the other hand, the only common zero of H_1 and H_2 is clearly the point $(0,0)$. Moreover, $[(\partial/\partial Z_1)H_1](0,0) = 1$, $[(\partial/\partial z_2)H_1](0,0) = 1$. Thus, in order to construct our example, we need only produce one function $F \varepsilon \boldsymbol{H}'$ which vanishes at $(0,0)$ but does not belong to I. We claim that such a function is $F(z_1, z_2) = z_1$.

To see that $z_1 \not\varepsilon I$, denote by I^* the (algebraic) ideal generated by H_1 and H_2 in $\mathfrak{H}$. Then by a result of H. Cartan, I^* is closed in $\mathfrak{H}$ (see [1], p. 62, Theorem 11). On the other hand, by [2], p. 297, Proposition 4, the topology of $\boldsymbol{H}'$ is stronger than that induced by $\mathfrak{H}$. Thus, $I^* \supset I$. But it is clear that there do not even exist formal power series G_1, G_2 in two variables such that $H_1 G_1 + H_2 G_2 = F$. Thus, $F \not\varepsilon I^*$, so $F \not\varepsilon I$. This completes the proof of our assertion.

Remark 1. Let us consider the change of coordinates $Z_1 \to X_1 - X_2$, $Z_2 \to X_2$. Then

$$(3) \qquad H_1 \to H_1 = X_1, \qquad H_2 \to H_2 = X_1 X_2 - X_2^2, \qquad F \to F = X_1 - X_2.$$

Thus, by (5), $[(\partial/\partial X_2)H_1](0) = 0$, $[(\partial/\partial X_2)H_2](0) = 0$, but $[(\partial/\partial X_2)F](0) = 1 \neq 0$. (This gives another proof that $F \not\varepsilon I$.) Thus, though we have shown that the answer to problem 2 of [2], p. 306 (as amended above) is in the negative, it seems reasonable to expect that the answer to the following problem is in the affirmative:

Problem 2′. Let J be a closed ideal in $\boldsymbol{H}'$, and let $G \varepsilon \mathbf{H}'$. Suppose not only that conditions (1) (for J instead of I) imply (2) (for G instead of F) but that the similar property holds after an arbitrary affine change of coordinates. Then, is $G \varepsilon J$?

An affirmative answer to Problem 2′ would clearly imply that every variety is determined by its exponential polynomials; the converse is, however, not obvious. Problem 2′ will be discussed in detail in a later publication.

Remark 2. It is a simple matter to extend the above to the case of more than two complex variables. Moreover, the space $\mathfrak{H}$ can be replaced by essentially any space such that the Fourier transform of its dual is a space of entire functions. In particular, we may replace $\mathfrak{H}$ by the space $\mathcal{E}$ of Schwartz (see [3], [4]).

THE INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] H. Cartan, "Idéaux et modules de fonctions analytiques de n variables complexes," *Bulletin de la Société Mathématique de France*, vol. 78 (1950), pp. 28-64.

[2] L. Ehrenpreis, "Mean periodic functions, Part I, Varieties whose annihilator ideal is principal," *American Journal of Mathematics*, vol. 77 (1955), pp. 293-328.

[3] L. Schwartz, "Theorie générale des fonctions moyenne-périodiques," *Annals of Mathematics*, vol. 48 (1947), pp. 857-929.

[4] ———, *Théorie des Distributions*, vol. I and II, Paris, 1951-52.

[5] ———, "Analyse et synthèse harmonique dans les espaces de distributions," *Canadian Journal of Mathematics*, vol. 3 (1951), pp. 503-512.

ON THE REGULAR SOLUTIONS AT A POINT OF SINGULARITY OF A SYSTEM OF NON-LINEAR DIFFERENTIAL EQUATIONS.*

By Robert W. Bass.[1]

1. Introduction. A theorem of Lettenmeyer [3] (which is a generalization of a theorem of Perron [5], reproduced in Bieberbach's recent *Theorie der gewöhnlichen Differentialgleichungen auf funktionentheoretischer Grundlage* (1953), pp. 161-172) concerns the existence of solutions of a system of linear differential equations

$$z^{s_i}w_i'(z)=\sum_{j=1}^{n}a_{ij}(z)w_j, \qquad (i=1,\cdots,n), \tag{1}$$

where $w_i'=dw_i/dz$. In these equations s_i $(i=1,\cdots,n)$ are non-negative integers such that

$$s=s_1+s_2+\cdots+s_n<n, \tag{2}$$

and the $a_{ij}(z)$ $(i,j=1,\cdots,n)$ are regular in some neighborhood of $z=0$. Then ([3]) the system (1) has at least $n-s$ linearly independent solutions regular at the origin.

The object of the present paper is to generalize this result of Lettenmeyer to the non-linear system

$$z^{s_i}w_i'(z)=f_i(z,w_1,\cdots,w_n), \qquad (i=1,\cdots,n), \tag{3}$$

in such a manner as to include the classical Cauchy-Poincaré theorem as a special case. This problem was suggested to me by Professor Wintner.

To be precise, let $f_i(z,w_1,\cdots,w_n)$ be n functions of $n+1$ variables, regular on some neighborhood of $(z,w_1,\cdots,w_n)=(0,0,\cdots,0)$ and such that their power-series representations converge absolutely on the closed product set

$$\text{(3a)}\quad |z|\leqq r_1, \qquad \text{(3b)}\quad |w_i|\leqq r_2, \qquad (i=1,\cdots,n).$$

Suppose further that $f_i(z,0,0,\cdots,0)\equiv 0$, so that the f_i can be written in the form

$$f_i=\sum_{j=1}^{n}a_{ij}(z)w_j+g_i(z,w_1,\cdots,w_n), \qquad (i=1,\cdots,n), \tag{4}$$

---

* Received May 13, 1955.

[1] National Science Foundation Predoctoral Fellow.

where the g_i are power series which converge absolutely on (3a), (3b) and contain only terms of the second degree or higher in the w_i.

THEOREM 1. *Let* (2) *and* (4) *hold. Then the system of differential equations* (3) *admits of at least an* $(n-s)$*-parameter algebroid family of solutions regular at* $z=0$. *The solutions will have algebroid dependence on any additional parameters in which the* f_i *are given regular.* (The precise meaning of "k-parameter algebroid family" and "algebroid dependence" will be set forth in Section 2.)

The technique of the proof constitutes an application of Wintner's theory ([7]) of analytic functions of infinitely many variables and of infinite implicit systems. We first suppose that (3) has a regular solution of the form

$$(5) \qquad w_i(z)=\sum_{j=0}^{\infty} d_{ij}z^j, \qquad (i=1,\cdots,n).$$

Then insert (5) in (3) and compare coefficients to obtain the necessary conditions

$$(6\nu) \qquad (\nu-s_i+1)d_{i\,\nu-s_i+1}=f_{i\nu}\,(d_{j0},d_{j1},\cdots,d_{j\nu};j=1,\cdots,n),$$

$(i=1,\cdots,n;\nu=0,1,2,\cdots)$. Here $d_{ik}\equiv 0$ for $k<0$ by definition, and the $f_{i\nu}$ are polynomials in $d_{j1},d_{j2},\cdots,d_{j\nu}$ $(j=1,\cdots,n)$ with coefficients regular in d_{j0} $(j=1,\cdots,n)$.

In the classical case, $s_i=0$ $(i=1,\cdots,n)$ and the conditions (6ν) are recursive; the coefficients d_{ij} for $j=2,3,\cdots$ are known explicitly as regular functions of the n parameters $d_{10},\cdots,d_{n0}$. By Cauchy's majorant method it is then proved ([6]) that for every $r<r_1$, an r_3 can be taken so small that (5) converge for $|z|\leqq r$ and satisfy (3b) whenever $|d_{i0}|<r_3$ $(i=1,\cdots,n)$ holds.

In the general case, however, the conditions (6ν) are not recursive. Note that by (2) we have in general excluded the Fuchsian theory of regular singular points $(s_i=1,g_i\equiv 0;i=1,\cdots,n)$, but have not excluded those regular singular points for which some f_i contains a factor z. (Nor have we excluded those irregular singular points which satisfy the necessary conditions ([1], p. 417) for the existence of $n-s$ solutions "regular" in the sense of [1].)

Suppose now that some s_i is positive, so that the system (6ν) is nonrecursive.

When the f_i are linear in the w_j (i.e., $g_i\equiv 0$), the conditions (6ν) are linear. Consequently the theory of infinite matrices may be applied. By using the resolvent series (C. Neumann), Perron ([5]) and Lettenmeyer ([3]) showed that the system (6ν) ($\nu = 0, 1, \cdots$) admits of at least $n - s$ linearly independent solutions, each of which is such that the corresponding solution (5) of (3) converges for all $|z| \leqq r_1$.

In the general case, the conditions (6ν) are neither recursive nor linear. Our method of proof will be to effect a change of variables such that Wintner's fixed-point theorem for analytic (but not in general completely continuous) mappings in a separable Hilbert space is applicable. This theorem allows us to conclude that for a sufficiently large integer N the components of the vector representing the fixed point will be regular in the d_{ij} ($i = 1, \cdots, n$; $j = 1, \cdots, N$) and any parameters in which the f_i are given regular. (The algebroid dependence occurs only among the d_{ij} ($i = 1, \cdots, n; j = 1, \cdots, N$).) Since the (modified) coefficients of the power series (5) obtained in this way are components of a vector in the Hilbert space we can easily prove the convergence of the solution (5).

Our change of variables is

$$\xi_{ik} = k d_{ik} r^k, \tag{7}$$

($i = 1, \cdots, n; k = 1, 2, \cdots$), $\xi_{i0} = d_{i0}$, which differs slightly from that of Lettenmeyer: our proof of the boundedness of the matrix of the linear terms of the $f_{i\nu}$ in (6ν) roughly parallels his proof, but by using (7) we can conclude directly (without explicit estimates) that the non-linear terms do form a mapping in Hilbert space. In fact, explicit estimates of his type would imply that the power series (5) converge only in the open circle $|z| < r$; but (7) and Schwarz's inequality imply that the series (5) converge absolutely on the closed circle

$$|z| \leqq r. \tag{8}$$

Now (8) allows the Parseval relation for the g_i to be applied at $|z| = r$, which produces the desired result.

We shall also generalize Theorem 1 as follows:

THEOREM 2. *Let* $f_i(z, w_1, \cdots, w_n)$, ($i = 1, \cdots, n$), *be as in Theorem* 1. *Let* $b_{ik}(z)$, ($i, k = 1, \cdots, n$), *be regular on some neighborhood of* $z = 0$ *and convergent on* (3a); *furthermore, assume that* $\det \| b_{ik}(z) \| \not\equiv 0$, *and that* $\det \| b_{ik}(z) \|$ *has an s-fold zero at* $z = 0$ *with* $s < n$. *Then the system of differential equations*

$$\sum_{k=1}^{n} b_{ik}(z) w_k'(z) = f_i(z, w_1, \cdots, w_n), \tag{9}$$

$(i=1,\cdots,n)$, has at least an $(n-s)$-parameter algebroid family of solutions regular at $z=0$. The solutions will have algebroid dependence on any additional parameters in which the f_i are given regular.

2. Definition. We shall say that (3) has an *m-parameter algebroid family* of solutions regular at the origin when the following conditions are satisfied:

(i) For every positive $r<r_1$ there exist a positive $r_3<r_2$, $2m$ non-negative integers i_k, j_k, $(k=1,\cdots,m)$, a positive integer N and n^2N^2 (complex) constants p_{ijkl}, $(i,k=1,\cdots,n;j,l=1,\cdots,N)$, such that for any choice of complex numbers μ_k, $(k=1,\cdots,m)$, satisfying

$$(10) \qquad |\mu_k|<r_3, \qquad (k=1,\cdots,m),$$

there exists a power series solution (5) of (3), absolutely convergent on (8), satisfying (3b), and such that if we put

$$(10') \qquad \mu_{ij}=\sum_{k=1}^{n}\sum_{l=1}^{N}p_{ijkl}d_{kl}, \qquad (i=1,\cdots,n;j=1,\cdots,N).$$

then

(ii) the transformation (10′) is non-singular, and

(iii) *all the initial conditions $d_{ij}\equiv w_i^{(j)}(0)$, $(i=1,\cdots,n;j=1,2,\cdots)$ are algebroid functions of the m parameters μ_k, where*

$$(11) \qquad \mu_k=\mu_{i_kj_k}, \qquad (k=1,\cdots,m).$$

If the f_i contain q additional parameters, then the μ_{ij} of (10′) depend linearly on them (as well as the d_{ik}), which by (11) implies the algebroid dependence of the solution on these $m+q$ parameters.

3. Proof of Theorem 1. Substitute (5) into $g_i(z,w_1,\cdots,w_n)$ and expand formally,

$$(12) \qquad g_i=\sum_{\nu=0}^{\infty}g_{i\nu}(d_{kl})z^{\nu} \qquad (i=1,\cdots,n),$$

to define the polynomials $g_{i\nu}=g_{i\nu}(d_{k1},\cdots,d_{k\nu};k=1,\cdots,n)$ whose coefficients are uniformly convergent power series in the d_{k0} $(k=1,\cdots,n)$. Also, let

$$(13) \qquad a_{ij}(z)=\sum_{k=0}^{\infty}a_{ijk}z^k, \qquad (i,j=1,\cdots,n).$$

Let M_1 be the largest maximum modulus of the g_i on (3a), (3b), and M_2 the largest for the a_{ij} on (3a). Let r be fixed such that $0 < r < r_1$. By the remark following (8), whenever (5) converges on (8) and satisfies (3b) there,

$$\sum_{k=0}^{\infty} | g_{ik} |^2 r^{2k} \leqq M_1^2, \qquad (i = 1, \cdots, n). \tag{14}$$

Since the g_i are, by (4), power series containing no terms of degree lower than the second, replacing the w_k in (3) by $\mu\, {}^1w_k$, $(k = 1, \cdots, n)$, replaces g_i in (12) by $(\mu^2)\,{}^1g_i(d_{kl}, \mu)$, where ${}^1g_i(d_{kl}, \mu)$ is a power series in μ. If this change of scale is made, (12) and (13) inserted in (16ν), and the result divided by μ, we have the equivalent conditions

$$(\nu - s_i + 1)\, d_{i\,\nu - s_i + 1} = \sum_{j=1}^{n} \sum_{\lambda=0}^{\nu} a_{ij\,\nu-\lambda}\, d_{j\lambda} + \mu\, {}^1g_i(d_{kl}, \mu). \tag{$15i\nu$}$$

Split off the $nN + s$ equations ($15i\nu$) for $\nu = 0, 1, \cdots, N + s_i - 1$, $(i = 1, \cdots, n)$, and call them (15a); call the remaining infinite system of equations (15b).

Now when (7) is used, ($15i\nu$) become

$$\xi_{i\,\nu - s_i + 1} = F_{i\nu}(\xi_{kl};\, l = 0, \cdots, \nu) + \mu G_{i\nu}, \tag{$15i\nu'$}$$

where (letting $1/\lambda = 1$ at $\lambda = 0$)

$$F_{i\nu} = r/r^{s_i} \sum_{j=1}^{n} \sum_{\lambda=0}^{\nu} (a_{ij\,\nu-\lambda} r^{\nu-\lambda}/\lambda)\, \xi_{j\lambda}; \tag{16a}$$

$$G_{i\nu} = (r/r^{s_i})\,[{}^1g_{i\nu}(d_{kl}, \mu)\, r^{\nu}] = (r/r^{s_i})\,[{}^1g_{i\nu}(\xi_{kl}/kr^k, \mu)\, r^{\nu}]. \tag{16b}$$

To consider (15b′) put $\nu - s_i = \alpha$ in ($15i\nu'$), (16a), and (16b), and require $\alpha \geqq N$ (N is to be fixed later). We then have

$$\xi_{1\,\alpha+1} = F_{i\,\alpha+s_i} + \mu G_{i\,\alpha+s_i}(\mu), \qquad (\alpha \geqq N). \tag{15b′}$$

Let r_3 be such that $0 < r_3 < 2r_3 < \min(r_1, r_2)$ and require that

$$| \xi_{io} | < 2r_3 < \min(r_1, r_2). \tag{17}$$

Suppose also (subject to later verification) that (15a′) can be solved in such a manner that

$$| \xi_{ik} | < 2\epsilon, \qquad (k = 1, \cdots, N), \tag{18}$$

where $0 < \epsilon \leqq r_3$. Since the right hand sides of (15b′) contain no ξ_{ij} with j greater than $\alpha + s_i$, the functions $F_{i\,\alpha+s_i}$ and $G_{i\,\alpha+s_i}$ are "analytic functions of the infinitely many variables $\xi_{i\,N+1}, \xi_{i\,N+2}, \cdots$ $(i=1,\cdots,n)$" in the sense of [7]. The remaining ξ_{ik} $(k=0,1,\cdots,N)$ are for the present to be regarded as parameters. Note that by (16a), (16b), and the remark following (12), the F_{ik} and G_{ik} converge uniformly in these parameters when the parameters are restricted by (17) and (18).

Let $\boldsymbol{M}_{2r_3}$ and $\boldsymbol{N}_{2\epsilon}$ be the regions (17) and (18), and define $\boldsymbol{M}_{r_3}$ and $\boldsymbol{N}_{\epsilon}$ similarly. Let $\boldsymbol{k}_{\Delta}$ represent the open disk $|\mu| < \Delta$ and $\boldsymbol{K}_{\delta}$ the sphere

$$\sum_{i=1}^{n}\sum_{k=N+1}^{\infty} |\xi_{ik}|^2 < \delta^2. \tag{19}$$

Fix δ such that

$$\delta < 3^{\frac{1}{2}}(r_2 - 2r_3)/\pi; \tag{20}$$

note that decreasing r_3 (which must be done later) does not disturb the inequality (20). At first suppose merely that $\epsilon \leqq r_3 < \delta$, so that whenever (17), (18), and (19) hold, the inequality

$$|\xi_{ik}| < 2\delta \tag{21}$$

holds for all i, k.

If we can show that for some fixed Δ, say $\Delta_1 > 0$, there exist numbers ω_N and ω such that, on the product set $\boldsymbol{K}_{\delta} \times \boldsymbol{M}_{2r_3} \times \boldsymbol{N}_{2\epsilon}$,

$$\sum_{i=1}^{n}\sum_{k=N}^{\infty} |F_{i\,k+s_i}|^2 < \omega_N, \tag{22}$$

$$\sum_{i=1}^{n}\sum_{k=N}^{\infty} |G_{i\,k+s_i}|^2 < \omega, \tag{23}$$

and that (keeping δ fixed) $\omega_N \to 0$ as $N \to \infty$, then by Theorem IIa of [7] there exist an N and a $\Delta < \Delta_1$ such that (15b) has exactly one solution $\xi_{kj}(\mu)$, analytic on $\boldsymbol{k}_{\Delta}$ and satisfying (19), for each point of $\boldsymbol{M}_{2r_3} \times \boldsymbol{N}_{2\epsilon}$. (This is true because by letting Δ be sufficiently small, not only ω_N but also $|\mu|^2\omega$ can be made sufficiently small in relation to the fixed δ.) This solution ξ_{kj} $(k=1,\cdots,n;\ j=N+1, N+2,\cdots)$ is analytic in all parameters on $\boldsymbol{k}_{\Delta} \times \boldsymbol{M}_{r_3} \times \boldsymbol{N}_{\epsilon}$; it is also analytic (for a halved radius of convergence) in any other parameters in which the f_i (and so the F's and G's) are given analytic.

We first consider the validity of (22). By (16a) and (21),

$$|F_{i\,\alpha+s_i}| < 2r\delta/r^{s_i}\sum_{j=1}^{n}\sum_{\lambda=0}^{\alpha+s_i} |a_{ij\,\alpha+s_i-\lambda}|\, r^{\alpha+s_i-\lambda}/\lambda. \tag{24}$$

Now by the choice of r, $\theta = r/r_1 < 1$. We next use Cauchy's coefficient estimate for (13) to establish that

$$|F_{i\,\alpha+s_i}| < 2nr\delta M_2/r^{s_i}\sum_{\lambda=0}^{\alpha+s_i}\theta^{\alpha+s_i-\lambda}/\lambda. \tag{25}$$

By splitting the sum in (25), using Schwarz's inequality on the first half and majorization by a geometric series on the second, we find

$$(26)\qquad \sum_{\lambda=0}^{\alpha+s_i} \theta^{\alpha+s_i-\lambda}/\lambda < c_1\theta^{(\alpha+s_i)/2} + 2c_2/(\alpha + s_1 - 1),$$

where $c_1 = [(1+\pi^2/6)/(1-\theta)]^{\frac{1}{2}}$ and $c_2 = 1/(1-\theta)$. By (25), (26), and Minkowski's inequality, we can take for ω_N in (22)

$$(27)\qquad \omega_N = n(2nr\delta M_2/r^{s_i})^2\Omega_N,$$

where $\Omega_N = [(c_1\theta^{N/2}/(1-\theta^{\frac{1}{2}})^{\frac{1}{2}}) + 2c_2(1/(N-1)^2 + \cdot\cdot\cdot)^{\frac{1}{2}}]^2$. Consequently, we can choose N so large that ω_N is as small as required.

We have assumed $\epsilon \leqq r_3 < \delta$. Now that N is fixed, we may further diminish r_3 if necessary, in order to ensure that

$$(28)\qquad 4nN\epsilon^2 \leqq 4nNr_3^2 < \delta.$$

We shall next prove (23), under the hypotheses (17), (18), (19). By (28), the hypotheses (18), (19) imply

$$(29)\qquad \sum_{j=1}^{n}\sum_{k=1}^{\infty} |\xi_{jk}|^2 < 2\delta^2.$$

This holds a fortiori if the first summation sign is removed. Hence by Schwarz's inequality and the fact that $\sum 1/k^2 = \pi^2/6$,

$$(30)\qquad \sum_{k=1}^{\infty} |d_{jk}|\, r^k = \sum_{k=1}^{\infty} |\xi_{jk}|/k \leqq \pi\delta/3^{\frac{1}{2}},$$

which proves that the series (5) converge absolutely on (8). Furthermore, by (17), (20), and (30),

$$(31)\qquad |w_i(z)| \leqq |d_{i0}| + \left|\sum_{k=1}^{\infty} d_{ik}z^k\right| < r_2.$$

In proving (23), we need only (by the maximum modulus theorem for the partial sums of (23), as stated in Littlewood [2], p. 238) consider the functions $g_{i\nu}(\mu)$ of (12), (14), and (16b) at $|\mu| = 1$, as will now be shown. Equation (31) (which continues to hold when the d_{ik} are each multiplied by $e^{i\phi}, \phi > 0$) shows that for any choice of the coefficients of (5) satisfying (17), (18), and, above all, (19), the power series (5) can be inserted in the g_i and expanded as in (12). Then (14) holds, and continues to hold, if the first $N + s_i - 1$ terms are omitted. Comparing this with (16b), we see that on $|\mu| = 1$ we may put

$$(32)\qquad \omega = r^2M_1{}^2\sum_{j=1}^{n} r^{-2s_j}.$$

But then by the argument of [2], (23) holds for all $|\mu| \leqq 1$.

Thus we have not only shown that the system of conditions (15b′) possesses solutions $\xi_{jk}=\xi_{jk}(\mu,\xi_{j0},\cdots,\xi_{jN};r)$, $(k=N+1,N+2,\cdots)$, analytic on $\boldsymbol{k}_\Delta\times\boldsymbol{M}_{r_3}\times\boldsymbol{N}_\epsilon$, but in (31) we have shown that the solutions (5) resulting from (15b′) converge on (8) and satisfy (3b), provided only that (15a′) and (15b′) can be satisfied simultaneously for values of ξ_{jk}, $(k=0,1,\cdots,N)$, lying in $\boldsymbol{M}_{r_3}\times\boldsymbol{N}_\epsilon$.

First, fix μ on $\boldsymbol{k}_\Delta$. Then the solutions of (15b′) just established can be rewritten by means of (7) in the original variables:

$$d_{j\,N+1}=\eta_{j1}(d_{j0},\cdots,d_{jN}),\cdots,d_{j\,N+\sigma-1}=\eta_{j\,\sigma-1}(d_{j0},\cdots,d_{jN}) \tag{33}$$

$$d_{j\,N+\rho}=\eta_{j\rho}(d_{j0},\cdots,d_{jN}), \qquad (\rho=\sigma,\sigma+1,\cdots) \tag{34}$$

where $\sigma=\max s_i$, $(i=1,\cdots,n)$. The functions η_{jk} are analytic in their $N+1$ variables. Since the solution (33)-(34) is unique and (3) admits the solution $w_i\equiv0$, $(i=1,\cdots,n)$, we know that $\eta_{jk}(0,0,\cdots,0)=0$ for all j,k.

The $\sum_{i=1}^{n}(N+s_i)=nN+s$ equations (15a′) have the form

$$\begin{gathered}0=f_{j0}(d_{j0}),\cdots,0=f_{j\,s_j-1}(d_{j0},\cdots,d_{j\,s_j-1})\\ d_{j1}=f_{js_j}(d_{j0},\cdots,d_{js_j}),\cdots,d_{jN}=f_{j\,N+s_j-1}(d_{j0},\cdots,d_{j\,N+s_j-1}).\end{gathered} \tag{35}$$

According to the definition of σ, none of the equations (35) contain variables which do not occur in the $n(\sigma-1)$ equations (33). Furthermore, $f_{jk}(0,\cdots,0)=0$, $(k=0,\cdots,N+s_j-1)$. Considering (33) and (35) as a single system, we then have $nN+s+n\sigma-n$ analytic implicit equations, vanishing at the origin, for the $n(N+\sigma)$ unknowns $d_{j0},\cdots,d_{j\,N+\sigma-1}$, $(j=1,\cdots,n)$. Consequently, by a theorem of Weierstrass (as stated in [4], p. 132), at least $n(N+\sigma)-(nN+s+n\sigma-n)=n-s$ linear combinations of the unknowns can be left as parameters, say the μ_k of (11), with $m\geqq n-s$; and the remaining unknowns can be found as algebroid functions of them. Moreover, the transformation (10′) is non-singular, and since these algebroid functions disappear at $\mu_k=0$, $(k=1,\cdots,m)$, the equations (17) and (18) can be satisfied (with r_3 and ϵ replaced by $\frac{1}{2}r_3$ and $\frac{1}{2}\epsilon$) by decreasing the value of r_3 which appears in (10).

4. Proof of Theorem 2. Let $c_{ik}(z)$, $d_{ik}(z)$, $(i,k=1,\cdots,n)$, be power series regular near $z=0$ and such that $\det\|c_{ik}\|$, $\det\|d_{ik}\|$ have no zeros in some neighborhood of the origin. In (9) put $w_k=\sum_{j=1}^{n}c_{ik}(z)v_k$; then (9) becomes a system (9′) of the same type for v_k. If f_i is given by (4), then

b_{ik} is replaced by $\sum_{l=1}^{n} b_{il}c_{lk}$ and a_{ik} replaced by $\sum_{j=1}^{n} (a_{ij}b_{jk} - b_{ij}c_{jk}')$. The new g_i are of the second degree or higher in the v_k. Now multiply the k-th equation of (9′) by $d_{ik}(z)$ and sum over k $(k = 1, \cdots, n)$; the result is a system (9″) of the form (9) again, but with b_{ik} replaced by $e_{ik} = \sum_{j=1}^{n} \sum_{l=1}^{n} d_{ij}(z) b_{jl}(z) c_{lk}(z)$. Lettenmeyer [3] has shown that $d_{ij}(z)$ and $c_{lk}(z)$ of the type mentioned can be so chosen that $e_{ik} = z^{s_i}\delta_{ik}$ $(\delta_{ik} = 0, i \neq k;\ = 1,\ i = k)$, where s_i are integers satisfying (2). Theorem 2 then follows as an immediate corollary of Theorem 1.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] E. L. Ince, *Ordinary differential equations*, 1927, p. 417.

[2] J. E. Littlewood, *Lectures on the theory of functions*, 1944, p. 238.

[3] F. Lettenmeyer, "Über die an einer Unbestimmtheitsstelle regulären Lösungen eines Systems homogener linearer Differentialgleichungen," *Sitzungsberichte München* (1926), pp. 287-307.

[4] W. F. Osgood, *Lehrbuch der Funktionentheorie*, 2nd ed. (1929), vol. II_1, p. 132.

[5] O. Perron, "Über diejenigen Integrale linearer Differentialgleichungen welche sich an einer Unbestimmtheitsstelle bestimmt verhalten," *Mathematische Annalen*, vol. 70 (1911), pp. 1-32.

[6] H. Poincaré, *Les méthodes nouvelles de la mécanique céleste*, vol. I (1892), p. 55.

[7] A. Wintner, "Zur Analyse im Hilbertschen Raume," *Mathematische Zeitschrift*, vol. 28 (1928), pp. 457-470.

REPRESENTATIONS OF SEMISIMPLE LIE GROUPS IV.*

By HARISH-CHANDRA.

1. **Introduction.** In an earlier paper of this series, I have given a method of constructing all irreducible quasi-simple representations of a connected semisimple group G (at least in case G has a finite-dimensional faithful representation) up to infinitesimal equivalence (see Theorem 4 of [5(d)]. However this method has one serious drawback. It can happen very well that the representation so obtained is infinitesimally unitary without being equivalent to a unitary representation. In order to make it unitary one has to introduce a new scalar product in a dense subspace of the representation space in some more or less artificial fashion and then complete this subspace with respect to the new norm. It would of course be much better if the unitary representations could be obtained in a more natural manner especially since it is often a rather difficult matter to decide whether a given representation is infinitesimally unitary or not. Suppose, for a moment, that G is a linear group and A is a Cartan subgroup of G. Then A can be written as a direct product of two subgroups A_- and A_+ of which the former is compact while the latter is isomorphic to a vector space. We say that A is incompact if A_+ has the maximum possible dimension. Now in the above-mentioned construction it was quite essential that one should start from an incompact A. However the study of simple examples (such as that of the 2×2 real unimodular group [1, 5(e)]) is enough to convince one of the necessity of searching for alternative 'natural' methods of construction which start out from other types of Cartan subgroups. The work of Gelfand and Graev [4] on the Plancherel formula of the $n \times n$ real unimodular group points to the same conclusion. In a recent note [5(f)] I have given such a 'natural' mode of construction based on the concept of induced representations (see Mackey [6]), which is applicable to all Cartan subgroups. By a simple argument the main problem can be reduced to the case when A is compact and therefore it is this particular case which requires intensive study. The object of the present paper is to deal with the algebraic aspects of this case, the function-theoretic aspects being left to a subsequent paper. Assuming that G is simple and its first Betti number is 1, we shall

---

* Received July 8, 1955.

give a quite general method of constructing a series of irreducible unitary representations of G parameterised by the characters of A. However if the first Betti number is zero (and G is not compact) this method does not give anything at all [5(g)]. (It seems likely that this is not due to a fault of the method but rather reflects an essential difference in the two cases with regard to representation theory. It is noteworthy that a similar difference has been observed by E. Cartan [3(d)] in his theory of bounded homogeneous symmetric domains. There also exactly those simple groups come into play whose first Betti number is one.) The representations obtained in this manner bear a very close resemblance, in their algebraic structure, to finite-dimensional representations. As we shall see in another paper, they are also somewhat remarkable in so far as their matrix coefficients are square-integrable on the group [5(f), (g)]. In the last section we obtain a formula (see Corollary to Lemma 21) for the multiplicities of the various weights of these representations.

Some of the results of this paper have been announced in a short note [5(g)].

2. Preliminary lemmas. Let R be the field of real numbers and V a vector space over R of finite dimension. We shall say that V is ordered if there is given an order in V satisfying the following condition. If u, v are two elements in V and a is a positive real number then $u+v>0$ and $au>0$ whenever $u>0$, $v>0$. An element $u \varepsilon V$ will be called positive (with respect to the given order) if $u>0$. Let $\mathfrak{F}$ be the space of all linear functions on V. Then there are two simple ways of introducing an order in $\mathfrak{F}$ which depend on the choice of a base either in V or in $\mathfrak{F}$. Let $H_1, \cdots, H_l$ be a base for V. We define a function $\lambda \varepsilon \mathfrak{F}$ to be positive if $\lambda \neq 0$ and $\lambda(H_i)>0$, i being the least index such that $\lambda(H_i) \neq 0$. This is called the lexicographic order in $\mathfrak{F}$ corresponding to the base $(H_1, \cdots, H_l)$. Similarly let $(\lambda_1, \cdots, \lambda_l)$ be a base for $\mathfrak{F}$. Then every $\lambda \varepsilon \mathfrak{F}$ can be written as $\lambda = c_1\lambda_1 + \cdots + c_l\lambda_l$ $(c_j \varepsilon R)$. We now define λ to be positive if $\lambda \neq 0$ and $c_i > 0$ where i is the least index such that $c_i \neq 0$. The order so defined is called the lexicographic order in $\mathfrak{F}$ corresponding to the base $(\lambda_1, \cdots, \lambda_l)$.

Let $\mathfrak{g}$ be a semisimple Lie algebra over the field C of complex numbers and let $\mathfrak{h}$ be a Cartan[1] subalgebra of $\mathfrak{g}$. Let $\mathfrak{h}_R$ denote the subset of those elements of $\mathfrak{h}$ at which every root α (of $\mathfrak{g}$ with respect to $\mathfrak{h}$) takes a real value. Then $\mathfrak{h}_R$ is a vector space over R and if $H_1, \cdots, H_l$ is a base for $\mathfrak{h}_R$

---

[1] We use the standard terminology of the theory of semisimple Lie algebras (see for example [5(b)]).

over R, it is also a base for $\mathfrak{h}$ over C. We say that a linear function λ on $\mathfrak{h}$ is real if $\lambda(H)$ is real for $H \varepsilon \mathfrak{h}_R$. Then the set $\mathfrak{F}_R$ of all such real functions may be identified in the obvious way with the space of all linear functions on $\mathfrak{h}_R$. Hence corresponding to the base $(H_1, \cdots, H_l)$ we get a lexicographic order in $\mathfrak{F}_R$. Similarly if $\alpha_1, \cdots, \alpha_l$ are l linearly independent roots they form a base for $\mathfrak{F}_R$ and we obtain the corresponding lexicographic order in $\mathfrak{F}_R$.

Let π be a representation of $\mathfrak{g}$ on a vector space V over C. A linear function Λ on $\mathfrak{h}$ is called a weight of π if there exists an element $\psi \neq 0$ in V such that $\pi(H)\psi = \Lambda(H)\psi$ for all $H \varepsilon \mathfrak{h}$. A weight Λ is called extreme if it is impossible to find a root α such that both $\Lambda + \alpha$ and $\Lambda - \alpha$ are also weights of π. It is well known (see Weyl [7]) that if $\dim V$ is finite every weight Λ is real.

For each root α choose an element $X_\alpha \neq 0$ in $\mathfrak{g}$ such that $[H, X_\alpha] = \alpha(H)X_\alpha$ for all $H \varepsilon \mathfrak{h}$. Then $H_\alpha = [X_\alpha, X_{-\alpha}]$ lies in $\mathfrak{h}$ and we may suppose that X_α is so chosen that $\alpha(H_\alpha) = 2$. Let $\mathfrak{B}$ be the universal enveloping algebra of $\mathfrak{g}$. We identify representations of $\mathfrak{g}$ with those of $\mathfrak{B}$ in the usual way.

LEMMA 1. *Let* π *be a representation of* $\mathfrak{g}$ *on* V *and let* $\psi \neq 0$ *be an element belonging*[1] *to the weight* Λ. *Suppose* α *is a root such that* $\pi(X_\alpha^r)\psi = \pi(X_{-\alpha}^r)\psi = 0$ *for some integer* $r \geqq 1$. *Then* $\Lambda(H_\alpha)$ *is an integer and* $\Lambda - k\alpha$ *is a weight of* π *for all integers* k *such that*

$$\min(0, \Lambda(H_\alpha)) \leqq k \leqq \max(0, \Lambda(H_\alpha)).$$

In particular $\Lambda - \Lambda(H_\alpha)\alpha$ *is a weight of* π. *Moreover if* $\pi(X_\alpha)\psi = 0$, $\Lambda(H_\alpha)$ *is equal to the smallest integer* $r \geqq 0$ *such that* $\pi(X_{-\alpha}^{r+1})\psi = 0$.

The proof of the first two statements is substantially the same as that of Lemma 11 of [5(b)]. Now suppose $\pi(X_\alpha)\psi = 0$. Let s be the least non-negative integer such that $\pi(X_{-\alpha}^{s+1})\psi = 0$. Then if $\psi_{-1} = 0$, $\psi_0 = \psi$ and $\psi_k = \pi(X_{-\alpha}^{k})\psi$ $(k \geqq 1)$ the argument of Lemma 11 of [5(b)] shows that

$$\pi(X_\alpha)\psi_k = k(\Lambda(H_\alpha) - k + 1)\psi_{k-1} \qquad (k \geqq 0).$$

Putting $k = s + 1$ we get $\Lambda(H_\alpha) = s$ and this proves the last part of assertion.

Let λ be a linear function on $\mathfrak{h}$. We say that an element $b \varepsilon \mathfrak{B}$ is of rank λ if $[H, b] = \lambda(H)b$ for all $H \varepsilon \mathfrak{h}$. Elements belonging to different ranks are linearly independent and if $\mathfrak{B}_\lambda$ is the set of all elements in $\mathfrak{B}$ of rank λ, $\mathfrak{B}$ is the direct sum of $\mathfrak{B}_\lambda$ for all λ [5(b), Lemma 2]. By the component of z of rank λ $(z \varepsilon \mathfrak{B})$ we mean the component of z in $\mathfrak{B}_\lambda$ in this direct sum decomposition.

Now suppose $\mathfrak{F}_R$ has been ordered in some way.

LEMMA 2. *Let π be a representation of $\mathfrak{B}$ on a vector space V. Suppose Λ is a linear function on $\mathfrak{h}$ and $\psi \neq 0$ a vector in V such that:*

(1) $\pi(X_\alpha)\psi = 0$ *for every positive root* α;

(2) $\pi(H)\psi = \Lambda(H)\psi$ *for all* $H \varepsilon \mathfrak{h}$;

(3) $V = \pi(\mathfrak{B})\psi$.

Then if $\alpha_1, \cdots, \alpha_r$ are all the distinct positive roots of $\mathfrak{g}$ every weight of π is of the form $\Lambda - (m_1\alpha_1 + \cdots + m_r\alpha_r)$ where $m_1, \cdots, m_r$ are nonnegative integers. Moreover if for each α_i we can find an integer $n_i \geqq 0$ such that $\pi(X_{-\alpha_i}{}^{n_i+1})\psi = 0$ then $\dim V$ *is finite and π is irreducible.*

Since $\mathfrak{g}$ is spanned by $(X_{\alpha_1}, \cdots, X_{\alpha_r}, \mathfrak{h}, X_{-\alpha_1}, \cdots, X_{-\alpha_r})$ it follows from Lemma 1 of [5(a)] and our assumptions above that V is spanned by elements of the form $\pi(X_{-\alpha_1}{}^{m_1} X_{-\alpha_2}{}^{m_2} \cdots X_{-\alpha_r}{}^{m_r})\psi$. In view of Lemma 10 of [5(b)] our first assertion now follows immediately. Moreover it is clear that $C\psi$ is the space of all elements in V which belong to the weight Λ.

Now suppose $\pi(X_{-\alpha_i}{}^{n_i+1})\psi = 0$, $1 \leqq i \leqq r$. Then we claim that V is spanned by $\pi(X_{-\alpha_1}{}^{\nu_1} \cdots X_{-\alpha_r}{}^{\nu_r})\psi$, $(0 \leqq \nu_i \leqq n_i, 1 \leqq i \leqq r)$. For let U be the space spanned by these elements. It would be sufficient to show that $\pi(X_{-\alpha_1}{}^{m_1} \cdots X_{-\alpha_r}{}^{m_r})\psi \varepsilon U$ for all $m_1, \cdots, m_r \geqq 0$. We shall do this by induction on $M = m_1 + \cdots + m_r$. If $M = 0$ the statement is obviously true and so we may suppose $M > 0$. Also we may assume that $m_i > n_i$ for some i. Let j be the largest index such that $m_j > n_j$. If $j = r$, $\pi(X_{-\alpha_r}{}^{m_r})\psi = 0$ and so our statement is again true. Hence suppose $j < r$. Put

$$z = X_{-\alpha_1}{}^{m_1} \cdots X_{-\alpha_{j-1}}{}^{m_{j-1}}, \qquad w = X_{-\alpha_{j+1}}{}^{m_{j+1}} \cdots X_{-\alpha_r}{}^{m_r}$$

$(z = 1$ if $j = 1)$. Then $zX_{-\alpha_j}{}^{m_j}w = z[X_{-\alpha_j}{}^{m_j}, w] + zwX_{-\alpha_j}{}^{m_j}$ and since $m_j > n_j$, $\pi(X_{-\alpha_j}{}^{m_j})\psi = 0$. Therefore

$$\phi = \pi(X_{-\alpha_1}{}^{m_1} \cdots X_{-\alpha_r}{}^{m_r})\psi = \pi(z[X_{-\alpha_j}{}^{m_j}, w])\psi.$$

But $X_{-\alpha_1}, \cdots, X_{-\alpha_r}$ span a subalgebra of $\mathfrak{g}$ and therefore from Lemma 1 of [5(a)] $z[X_{-\alpha_j}{}^{m_j}, w]$ can be written as a linear combination of $X_{-\alpha_1}{}^{m_1'} \cdots X_{-\alpha_r}{}^{m_r'}$ with $m_1' + \cdots + m_r' < M$. Hence by our induction hypothesis $\phi \varepsilon U$. This proves that $U = V$ and therefore $\dim V$ is finite and V is fully reducible under π. Hence there exists a subspace of V' which is invariant and irreducible under $\pi(\mathfrak{B})$ and which contains a nonzero vector ψ' belonging to the weight Λ. But then, as we have seen above $\psi' = c\psi$ $(c \varepsilon C)$ and therefore $V' = \pi(\mathfrak{B})\psi' = V$. This proves that π is irreducible.

We shall need the following lemma due to E. Cartan [3(a)].

LEMMA 3 (Cartan). *Suppose π is an irreducible finite-dimensional representation of* $\mathfrak{g}$ *and Λ_0 is its highest weight. Then if W is the Weyl*[1] *group of* $\mathfrak{g}$ (*with respect to* $\mathfrak{h}$) $s\Lambda_0$ ($s \varepsilon W$) *are exactly all the extreme weights of π.*

We sketch the proof for the sake of completeness. It follows from the argument of Lemma 13 of [5(b)] that if Λ is an extreme weight of π so is also $s\Lambda$ for every $s \varepsilon W$. Since the highest weight is obviously an extreme weight, $s\Lambda_0$ ($s \varepsilon W$) are all extreme weights. Conversely suppose Λ is an extreme weight of π. Let Λ_1 be the highest one among $s\Lambda$ ($s \varepsilon W$). Then Λ_1 is also an extreme weight and it would be enough to prove that $\Lambda_1 = \Lambda_0$. Let α be a positive root and s_α the corresponding Weyl reflexion. Then $s_\alpha\Lambda_1 = \Lambda_1 - \Lambda_1(H_\alpha)\alpha \leqq \Lambda_1$. Hence $\Lambda_1(H_\alpha) \geqq 0$. We claim that $\Lambda_1 + \alpha$ is not a weight of π. For if it were we could conclude from Lemma 1 (since $\pi(X_\alpha)$ and $\pi(X_{-\alpha})$ are both nilpotent) that $\Lambda_1 + \alpha - \nu\alpha$ is a weight for $0 \leqq \nu \leqq \Lambda_1(H_\alpha) + 2$. This would imply however that $\Lambda_1 + \alpha$ and $\Lambda_1 - \alpha$ are both weights, contradicting the fact that Λ_1 is extreme. Let $\psi \neq 0$ be an element in V belonging to the weight Λ_1. Since $\Lambda_1 + \alpha$ is not a weight, $\pi(X_\alpha)\psi = 0$ and it follows from Lemma 2 that Λ_1 is the highest weight of π. Therefore $\Lambda_1 = \Lambda_0$ and Lemma 3 is proved.

A set $(\alpha_1, \cdots, \alpha_l)$ of roots of $\mathfrak{g}$ is called a fundamental system if it is a linearly independent set and if every root α can be written in the form $\alpha = d_1\alpha_1 + \cdots + d_l\alpha_l$ where d_i are integers which are either all nonnegative or all nonpositive.

LEMMA[2] 4. *Let* $\mathfrak{m}$ *be a subalgebra of* $\mathfrak{g}$ *such that* $\mathfrak{h} \subset \mathfrak{m}$ *and for each root α either X_α or $X_{-\alpha}$ is in* $\mathfrak{m}$. *Then it is possible to choose a fundamental system $(\alpha_1, \cdots, \alpha_l)$ of roots such that $X_{\alpha_i} \varepsilon \mathfrak{m}$, $1 \leqq i \leqq l$.*

Let $\mathfrak{r}$ be the radical of $\mathfrak{m}$ and let $X \to X^*$ denote the natural mapping of $\mathfrak{m}$ on $\mathfrak{m}^* = \mathfrak{m}/\mathfrak{r}$. Then $\mathfrak{m}^*$ is a semisimple Lie algebra and $\mathfrak{h} = (\mathfrak{h} + \mathfrak{r})/\mathfrak{r}$ is an abelian subalgebra of $\mathfrak{m}^*$. Let $X \to adX$ denote the adjoint representation of $\mathfrak{g}$. Then if $H \varepsilon \mathfrak{h}$, adH is a semisimple[3] linear transformation of $\mathfrak{g}$. Hence if ad^*H^* denotes the linear mapping of $\mathfrak{m}^*$ corresponding to H^* under the adjoint representation of $\mathfrak{m}^*$, it is clear that ad^*H^* is also semisimple. Therefore $\mathfrak{h}^*$ can be extended to a Cartan subalgebra $\mathfrak{h}_1{}^*$ of $\mathfrak{m}^*$. We introduce some lexicographic order among the roots of $\mathfrak{m}^*$ with respect to $\mathfrak{h}_1{}^*$. For every root γ^* of $\mathfrak{m}^*$ choose an element $X_{\gamma^*}{}^* \neq 0$ in $\mathfrak{m}^*$ such that $[H^*, X_{\gamma^*}{}^*] = \gamma^*(H^*)X_{\gamma^*}{}^*$ for all $H^* \varepsilon \mathfrak{h}_1{}^*$. Let $\mathfrak{r}_1{}^* = \mathfrak{h}_1{}^* + \sum_{\gamma^*>0} CX_{\gamma^*}$ and let Γ_1 be the complete inverse image of Γ_1^* in $\mathfrak{m}$. It is clear that Γ_1^* is a solvable subalgebra of $\mathfrak{m}^*$. Hence Γ_1 is a solvable subalgebra of $\mathfrak{m}$ containing $\mathfrak{h} + \Gamma$.

---

[2] A very similar result has been obtained independently by A. Borel.

[3] This means that $\mathfrak{g}$ is fully reducible under adH.

Now let β be any root of $\mathfrak{g}$. We claim that at least one of the two elements X_β or $X_{-\beta}$ is in Γ_1. In order to prove this we may clearly assume that $X_\beta \in \mathfrak{m}$ but $X_\beta \notin \Gamma_1$. Then $[H, X_\beta] = \beta(H)X_\beta$ $(H \in \mathfrak{h})$ and by going over to $\mathfrak{m}^*$ we get $[H^*, X_\beta^*] = \beta(H)X_\beta^*$ $(H \in \mathfrak{h})$. Since $X_\beta \notin \Gamma_1$, $X_\beta^* \neq 0$ and so we can choose a root γ^* of $\mathfrak{m}^*$ such that $\gamma^*(H^*) = \beta(H)$ for all $H \in \mathfrak{h}$. Define $\epsilon = 1$ or -1 according as γ^* is positive or negative and put $Y^* = X_{\epsilon\gamma^*}^* \in \Gamma_1^*$. Then if Y is any element in Γ_1 whose image in $\mathfrak{m}^*$ is Y^*, $Y \neq 0$ and

$$[H, Y] \equiv \epsilon\beta(H)Y \mod \Gamma \qquad (H \in \mathfrak{h}).$$

However Γ_1 is invariant and fully reducible under $ad\,\mathfrak{h}$. Therefore we can select an element $Z \in \Gamma_1$ $(Z \neq 0)$ such that

$$[H, Z] = \epsilon\beta(H)Z.$$

But then Z can differ from $X_{\epsilon\beta}$ only by a nonzero factor in C and therefore $X_{\epsilon\beta} \in \Gamma_1 \subset \mathfrak{m}$. This proves our assertion.

Put $B(X, Y) = sp(adXadY)$ $(X, Y \in \mathfrak{g})$. For any linear function λ on $\mathfrak{h}$ let H_λ' denote the (unique) element in $\mathfrak{h}$ such that $B(H_\lambda', H) = \lambda(H)$ for all $H \in \mathfrak{h}$. Define $\langle\lambda, \mu\rangle = B(H_\lambda', H_\mu') = \lambda(H_\mu')$ for any two linear functions λ, μ. Then it is well known [7] that if α is a root $H_\alpha' = c_\alpha H_\alpha$ where c_α is a *positive* real number. Let $\gamma_1, \cdots, \gamma_l$ be a fundamental system of roots of $\mathfrak{g}$. We introduce the lexicographic order (in $\mathfrak{F}_R$) corresponding to this system. Let Λ_0 be a linear function on $\mathfrak{h}$ such that $\Lambda_0(H_{\gamma_i})$ $1 \leqq i \leqq l$ are all positive integers. Now if α is a positive root $\alpha = d_1\gamma_1 + \cdots + d_l\gamma_l$ where d_i are nonnegative integers and therefore $H_\alpha' = d_1 H_{\gamma_1}' + \cdots + d_l H_{\gamma_l}'$. Hence

$$c_\alpha\Lambda_0(H_\alpha) = \Lambda_0(H_\alpha') = \sum_{i=1}^{l} d_i\Lambda_0(H_{\gamma_i}') = \sum_i d_i c_{\gamma_i}\Lambda_0(H_{\gamma_i})$$

and in view of the fact that $c_\alpha, c_{\gamma_i}, \Lambda_0(H_{\gamma_i})$ are all positive and not all d_i are zero we can conclude that $\langle\Lambda_0, \alpha\rangle$ and $\Lambda_0(H_\alpha)$ are both positive. Let π be an irreducible finite-dimensional representation of $\mathfrak{g}$ with the highest weight Λ_0. (We know from Theorem 1 of [5(b)] that such a representation exists). Since Γ_1 is a solvable algebra, by Lie's theorem we can choose an element $\psi_0 \neq 0$ in the representation space V and a linear function μ on Γ_1 such that $\pi(X)\psi_0 = \mu(X)\psi_0$ for all $X \in \Gamma_1$. Since $\mathfrak{h} \subseteq \Gamma_1$ it follows that μ coincides on $\mathfrak{h}$ with some weight Λ of π. Let β be a root such that $X_\beta \in \Gamma_1$.

Then ψ_0 is an eigenvalue of $\pi(X_\beta)$. But $\pi(X_\beta)$ and $\pi(X_{-\beta})$ are both nilpotent and therefore $\pi(X_\beta)\psi_0 = 0$ and $\pi(X_{-\beta}^r)\psi_0 = 0$ for r sufficiently large. Hence it follows from Lemma 1 that $\Lambda(H_\beta) \geqq 0$ and $\Lambda(H_\beta) \neq 0$ unless $\pi(X_{-\beta})\psi_0 = 0$. Let $(\beta_1, \cdots, \beta_s)$ be all the distinct roots β such that $X_\beta \varepsilon \Gamma_1$ and among these let $(\beta_1, \cdots, \beta_r)$ be all those for which $\Lambda(H_{\beta_i}) \neq 0$. Then $\Lambda(H_{\beta_i}) > 0$, $1 \leqq i \leqq r$ and in view of what we have proved above, $\mathfrak{g}$ is spanned by $\mathfrak{h}$, X_{β_i} and $X_{-\beta_i}$, $1 \leqq i \leqq s$. Hence it follows from Lemma 1 of [5(a)] that $\pi(\mathfrak{B})\psi_0 = V$ is spanned by elements of the form $\pi(X_{-\beta_1}^{m_1} \cdots X_{-\beta_s}^{m_s})\psi_0$ where $m_1, \cdots, m_s$ are nonnegative integers. Since $\pi(X_{-\beta_j})\psi_0 = 0$, $r < j \leqq s$, V is actually spanned by $\pi(X_{-\beta_1}^{m_1} \cdots X_{-\beta_r}^{m_r})\psi$. This proves that every weight of π must be of the form $\Lambda - (m_1\beta_1 + \cdots + m_r\beta_r)$, $(m_i \geqq 0)$. We now claim that Λ is an extreme weight of π. For otherwise let α be a root such that $\Lambda + \alpha$ and $\Lambda - \alpha$ are both weights. Then in view of the above result

$$\alpha = m_1\beta_1 + \cdots + m_r\beta_r, \qquad -\alpha = m_1'\beta_1 + \cdots + m_r'\beta_r$$

where $m_i, m_i' \geqq 0$. Hence $0 = (m_1 + m_1')\beta_1 + \cdots + (m_r + m_r')\beta_r$ and so

$$(m_1 + m_1')\langle\Lambda, \beta_1\rangle + (m_2 + m_2')\langle\Lambda, \beta_2\rangle + \cdots + (m_r + m_r')\langle\Lambda, \beta_r\rangle = 0.$$

But we have seen above that $\Lambda(H_{\beta_i}) > 0$ and therefore $\langle\Lambda, \beta_i\rangle > 0$, $i = 1, \cdots, r$. Hence the above equation implies that $m_i = m_i' = 0$, $1 \leqq i \leqq r$. This however is impossible since $\alpha \neq 0$. Therefore Λ is an extreme weight and we can conclude from Lemma 3 that $\Lambda = s\Lambda_0$ for some $s \varepsilon W$.

We recall that $\langle\Lambda_0, \alpha\rangle \neq 0$ for every root α. Hence $\langle\Lambda, \alpha\rangle = \langle\Lambda_0, s^{-1}\alpha\rangle \neq 0$ and therefore $\Lambda(H_\beta) > 0$ for every root β such that X_β lies in Γ_1. This shows in particular that for any root β exactly one of the two elements X_β and $X_{-\beta}$ is in Γ_1. Now put $\alpha_i = s\gamma_i$, $1 \leqq i \leqq l$. Then $(\alpha_1, \cdots, \alpha_l)$ is also a fundamental system of roots of $\mathfrak{g}$ and $\langle\Lambda, \alpha_i\rangle = \langle\Lambda_0, \gamma_i\rangle > 0$. Therefore $\Lambda(H_{\alpha_i}) > 0$ and hence $X_{\alpha_i} \varepsilon \Gamma_1 \subset \mathfrak{m}$. This proves the lemma.

COROLLARY 1. *Let P be a set of roots of* $\mathfrak{g}$ *(with respect to* $\mathfrak{h}$*) satisfying the following two conditions.*

(1) *For any root* β *exactly one of the two roots* β *and* $-\beta$ *lies in* P;

(2) *If* α *and* β *are in* P *and* $\alpha + \beta$ *is a root then* $\alpha + \beta \varepsilon P$.

Then it is possible to select roots $\alpha_i \varepsilon P$, $1 \leqq i \leqq l$ *such that* $(\alpha_1, \cdots, \alpha_l)$ *is a fundamental system of roots and this system is unique (apart from a permutation of its elements). Moreover* P *is exactly the set of positive roots in the lexicographic order corresponding to* $(\alpha_1, \cdots, \alpha_l)$.

It is clear that $\mathfrak{n} = \sum_{\beta \varepsilon P} CX_\beta$ is a subalgebra of $\mathfrak{g}$ and $\mathfrak{m} = \mathfrak{h} + \mathfrak{n}$ fulfils all the conditions of Lemma 4. Hence we can choose a fundamental system $(\alpha_1, \cdots, \alpha_l)$ such that $X_{\alpha_i} \varepsilon \mathfrak{m}$. It is obvious that $\alpha_i \varepsilon P$. Let β be any positive root in the lexicographic order corresponding to this system. Then X_β is contained in the subalgebra of $\mathfrak{g}$ generated by X_{α_i}, $1 \leqq i \leqq l$ (see Lemma 18 of [5(b)]). Hence it follows from the second condition of our hypothesis that $X_\beta \varepsilon \mathfrak{n}$ and therefore $\beta \varepsilon P$. So every positive root is in P and hence in view of condition (1) no negative root can lie in P. This proves that P is exactly the set of all positive roots. Finally it is obvious that if α is a root in P, it lies in the fundamental system $(\alpha_1, \cdots, \alpha_l)$ if and only if it cannot be written in the form $\alpha = \beta_1 + \cdots + \beta_r$ where $\beta_1, \cdots, \beta_r$ are all roots in P and $r \geqq 2$. This shows that the system $(\alpha_1, \cdots, \alpha_l)$ is unique apart from order.

COROLLARY 2. *Let P be the set of all positive roots under any given order in $\mathfrak{F}_R$. Then there exists a fundamental system $(\alpha_1, \cdots, \alpha_l)$ consisting of positive roots and this system is unique apart from a permutation. Moreover P is also the set of all positive roots in the lexicographic order corresponding to $(\alpha_1, \cdots, \alpha_l)$.*

For it is obvious that P satisfies the two conditions of Corollary 1.

3. Representations with an extreme vector.

DEFINITION. *Let π be a representation of $\mathfrak{B}$ on a vector space V. An element ψ in V will be called an extreme vector of π (with respect to $\mathfrak{h}$) if the following two conditions are fulfilled:*

(1) *$\psi \neq 0$ and there exists a linear function Λ on $\mathfrak{h}$ such that $\pi(H)\psi = \Lambda(H)\psi$ for all $H \varepsilon \mathfrak{h}$.*

(2) *For any root β at least one of the two elements $\pi(X_\beta)\psi$, $\pi(X_{-\beta})\psi$ is zero.*

Suppose ψ is an extreme vector of π. Let $\mathfrak{m}$ be the subalgebra of $\mathfrak{g}$ consisting of all elements $X \varepsilon \mathfrak{g}$ such that ψ is an eigenvector of $\pi(X)$. Then $\mathfrak{m}$ satisfies the conditions of Lemma 4. If β is a root such that $\pi(X_\beta)\psi \neq 0$, $\pi(X_\beta)\psi$ belongs to the weight $\Lambda + \beta$ and therefore $X_\beta \not\varepsilon \mathfrak{m}$. Hence if $X_\beta \varepsilon \mathfrak{m}$, $\pi(X_\beta)\psi = 0$. Therefore from Lemma 4 it is possible to choose a fundamental system of roots $(\alpha_1, \cdots, \alpha_l)$ such that $\pi(X_{\alpha_i})\psi = 0$, $i = 1, \cdots, l$. Then if we consider the lexicographic order corresponding to this system and set $\mathfrak{n} = \sum_{\beta > 0} CX_\beta$, $\mathfrak{n} \subset \mathfrak{m}$ (Lemma 18 of [5(b)] and therefore $\pi(X_\beta)\psi = 0$ for all $\beta > 0$. If $\beta_1, \cdots, \beta_r$ are all the distinct positive roots and we assume that $V = \pi(\mathfrak{B})\psi$, it follows from Lemma 2 that every weight of π is of the form $\Lambda-(m_1\beta_1+\cdots+m_r\beta_r)$, $(m_i\geqq 0)$ and therefore also of the form $\Lambda-(d_1\alpha_1+\cdots+d_l\alpha_l)$ where $d_1,\cdots,d_l$ are nonnegative integers.

Conversely suppose we are already given an order in $\mathfrak{F}_R$ and ψ is an extreme vector of the representation π. We say that ψ is a *positive* extreme vector if $\pi(X_\beta)\psi=0$ for all positive roots β. The above remarks show that any extreme vector is positive with respect to the lexicographic order corresponding to some fundamental system of roots.

LEMMA 5. *Let* Λ *be a linear function on* $\mathfrak{h}$. *Then there exists at most one maximal left ideal* $\mathfrak{M}$ *in* $\mathfrak{B}$ *such that*[4]

$$\mathfrak{M}\supset\sum_{\beta>0}\mathfrak{B}X_\beta+\sum_{\beta>0}\mathfrak{B}(H_\beta-\Lambda(H_\beta))$$

where β *runs over all positive roots.*

This is proved in exactly the same way as Lemma 1 of [5(b)].

COROLLARY. *Let* π_1, π_2 *be irreducible representations of* $\mathfrak{B}$ *on the vector spaces* V_1, V_2 *respectively. Suppose there exists an order in* $\mathfrak{F}_R$ *such that with respect to this order both* π_1 *and* π_2 *have positive extreme vectors belonging to the same weight* Λ. *Then* π_1, π_2 *are equivalent.*

Let ψ_i be a positive vector of π_i belonging to the weight Λ $(i=1,2)$. Let $\mathfrak{M}_i$ be the set of all $b\in\mathfrak{B}$ such that $\pi_i(b)\psi_i=0$. Since π_i is irreducible $\mathfrak{M}_i$ is a maximal left ideal in $\mathfrak{B}$ and π_i is equivalent to the natural representation (see [5(c), § 3]) of $\mathfrak{B}$ on $\mathfrak{B}/\mathfrak{M}_i$. But $\mathfrak{M}_1=\mathfrak{M}_2$ from Lemma 5 and so our assertion follows.

For reasons which we have explained briefly in Section 1 (see also [5(f)]), we shall now consider the following special case. Let $\mathfrak{g}_0$ be a semisimple Lie algebra over R. Henceforward I shall use the notation of my two previous papers [5(c), (d)]. Define $\mathfrak{k}_0$ and $\mathfrak{p}_0$ as in [5(c), § 2] and let $\mathfrak{h}_0$ be a maximal abelian subalgebra of $\mathfrak{k}_0$. *We shall assume that* $\mathfrak{h}_0$ *is also maximal abelian in* $\mathfrak{g}_0$. Complexify $\mathfrak{g}_0$ to $\mathfrak{g}$ and let $\mathfrak{h}$, $\mathfrak{p}$, $\mathfrak{k}$ denote the subspaces of $\mathfrak{g}$ spanned over C by $\mathfrak{h}_0$, $\mathfrak{p}_0$, $\mathfrak{k}_0$ respectively. Then $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$ and $[\mathfrak{h},\mathfrak{k}]\subset\mathfrak{k}$, $[\mathfrak{h},\mathfrak{p}]\subset\mathfrak{p}$. Therefore if β is any root (of $\mathfrak{g}$ with respect to $\mathfrak{h}$) it is clear that X_β lies either in $\mathfrak{k}$ or in $\mathfrak{p}$. We say that β is compact or noncompact according as $X_\alpha\in\mathfrak{k}$ or $X_\alpha\in\mathfrak{p}$. Two roots α and β are unlike if one of them is compact while the other is not. Otherwise they are like. Since $[\mathfrak{p},\mathfrak{p}]\subset\mathfrak{k}$, $[\mathfrak{k},\mathfrak{p}]\subset\mathfrak{p}$ and $[X_\alpha,X_{-\alpha}]=H_\alpha\in\mathfrak{h}\subset\mathfrak{k}$, it follows that α and $-\alpha$ are always like. If α,β and $\alpha+\beta$ are roots, it is known that

---

[4] That $\mathfrak{M}$ always exists follows from Lemma 16

$[X_\alpha, X_\beta] = cX_{\alpha+\beta}$ where $c \neq 0$ $(c \in C)$. Therefore $\alpha+\beta$ is compact or not according as α, β are like or unlike.

Let π be a representation of $\mathfrak{g}$ on a vector space V with an extreme vector ψ_0 (with respect to $\mathfrak{h}$) belonging to the weight Λ_0. Introduce a lexicographic order so that $\pi(X_\beta)\psi_0 = 0$ for every positive root β. Let P be the set of all positive roots and P_k and P_n respectively the subsets of P consisting of all compact and noncompact roots in P. Then it is clear that

$$\mathfrak{k} = \mathfrak{h} + \sum_{\alpha \in P_k} (CX_\alpha + CX_{-\alpha}), \qquad \mathfrak{p} = \sum_{\beta \in P_n} (CX_\beta + CX_{-\beta}).$$

In accordance with our earlier notation [5(c)] we denote by $\mathfrak{X}$ the subalgebra of $\mathfrak{B}$ generated by $(1, \mathfrak{k})$. Since we are ultimately interested in the irreducible representations of the simply connected group corresponding to the algebra $\mathfrak{g}_0$, *we shall now make the further assumption that* $\dim \pi(\mathfrak{X})\psi_0$ *is finite and* $V = \pi(\mathfrak{B})\psi_0$.

Since $\mathfrak{X}$ is the universal enveloping algebra of $\mathfrak{k}$, we can identify representations of $\mathfrak{k}$ with those of $\mathfrak{X}$. Let σ denote the representation of $\mathfrak{k}$ on $\mathfrak{g}$ given by $\sigma(X)Y = [X, Y]$ $(X \in \mathfrak{k}, Y \in \mathfrak{g})$.

DEFINITION. *A root β of $\mathfrak{g}$ will be called totally positive if there exists a subset P_β of P such that* $\sigma(\mathfrak{X})X_\beta = \sum_{\gamma \in P_\beta} CX_\gamma$.

It is clear that every totally positive root is positive. Later (in Section 4) we shall give another equivalent definition of totally positive roots (see the corollary to Lemma 12). Let P_+ be the set of all totally positive roots and Q the remaining set of positive roots. Our main object at present is to prove the following theorem.

THEOREM 1. *For every $\beta \in Q$, $\Lambda_0(H_\beta)$ is a nonnegative integer and it is the least integer $r \geqq 0$ such that* $\pi(X_{-\beta}^{r+1})\psi_0 = 0$.

Let Q' denote the set of all positive roots β such that $\pi(X_{-\beta}^{r})\psi_0 = 0$ for some $r \geqq 1$. Since $\pi(X_\beta)\psi_0 = 0$ for $\beta \in P$, it follows from Lemma 1 that in order to prove the above theorem it is enough to show that $Q \subset Q'$. For this we need some lemmas.

LEMMA 6. *Suppose $\beta \in Q'$. Then for any $\psi \in V$ we can choose an integer $r \geqq 1$ such that* $\pi(X_\beta^{r})\psi = \pi(X_{-\beta}^{r})\psi = 0$.

Let $\beta_1, \cdots, \beta_t$ be all the distinct positive roots. As we have already observed earlier $V = \pi(\mathfrak{B})\psi_0$ is then spanned by elements of the form $\pi(X_{-\beta_1}^{m_1} \cdots X_{-\beta_t}^{m_t})\psi_0$, $(m_j \geqq 0)$. Hence it is enough to prove our statement for $\psi = \pi(x)\psi_0$ where $x = X_{-\beta_1}^{m_1} \cdots X_{-\beta_t}^{m_t} \in \mathfrak{B}$. Now if $X \in \mathfrak{g}$,

$Xx = xX + [X, x]$. Therefore if we denote by adX the endomorphism $b \to [X, b]$ $(b \varepsilon \mathfrak{B})$ of $\mathfrak{B}$, it follows that (see pp. 40-41 of [5(b)])

$$X^p x = \sum_{0 \leqq q \leqq p} C_q{}^p \{(adX)^{p-q} x\} X^q$$

for any integer $p \geqq 0$. (Here $C_q{}^p = p!/(p-q)!q!$). However if $X = X_\beta$ or $X_{-\beta}$ the restriction of adX on $\mathfrak{g}$ is nilpotent and from this it follows easily that $(adX)x^m = 0$ if m is sufficiently large. Since $\beta \varepsilon Q'$, $\pi(X_\beta)\psi_0 = 0$ and we can choose an integer $s \geqq 1$ such that $\pi(X_{-\beta}{}^s)\psi_0 = 0$. Then if we put $p = m + s$ we get

$$X^{m+s} = \sum_{0 \leqq q \leqq m+s} C_q{}^{m+s}\{(adX)x^{m+s-q}\}X^q = \sum_{s < q \leqq m+s} C_q{}^{m+s}\{(adX)x^{m+s-q}\}X^q$$

since $(adX)x^m = 0$. On the other hand $\pi(X^q)\psi_0 = 0$ for $q \geqq s$ and therefore $\pi(X^{m+s})\psi = \pi(Xx^{m+s})\psi_0 = 0$. This proves our assertion.

COROLLARY. *Let W' be the subgroup of the Weyl group W generated by the Weyl reflexions s_β corresponding to $\beta \varepsilon Q'$. Then if Λ is a weight of π and $s \varepsilon W'$, $s\Lambda$ is also a weight of π.*

Suppose $s = s_{\beta_1} s_{\beta_2} \cdots s_{\beta_r}$ where $\beta_i \varepsilon Q'$. If $r = 1$ the result follows immediately from Lemmas 6 and 1. Hence assume $r \geqq 2$ and use induction on r. Then if $s' = s_{\beta_2} \cdots s_{\beta_r}$ it follows from our induction hypothesis that $s'\Lambda$ is a weight of π. Hence again by the above argument $s_{\beta_1}(s'\Lambda) = s\Lambda$ is also a weight of π.

LEMMA 7. *Let β be a positive root. Suppose $s\beta < 0$ for some $s \varepsilon W'$. Then $\beta \varepsilon Q'$.*

Let $\beta_1, \cdots, \beta_l$ be a fundamental system of positive roots (see Corollary 2 to Lemma 4). Then we know from Lemma 2 that every weight of π must be of the form $\Lambda_0 - (m_1\beta_1 + \cdots + m_l\beta_l)$ where m_i are nonnegative integers. Now suppose $\pi(X_{-\beta}{}^r)\psi_0 \neq 0$ for some positive integer. Then $\Lambda_0 - r\beta$ is a weight of π. Since $s \varepsilon W'$ it follows from the above corollary that both $s\Lambda_0$ and $s(\Lambda_0 - r\beta)$ are weights of π. Therefore

$$s\Lambda_0 = \Lambda_0 - (d_1\beta_1 + \cdots + d_l\beta_l) \qquad (d_i \geqq 0)$$

$$s\Lambda_0 - rs\beta = \Lambda_0 - (m_1\beta_1 + \cdots + m_l\beta_l) \qquad (m_i \geqq 0).$$

But $s\beta$ is a negative root and therefore $s\beta = -(e_1\beta_1 + \cdots + e_l\beta_l)$ $(e_i \geqq 0)$, e_i being integers which are not all zero. Hence

$$\begin{aligned} -r(e_1\beta_1 + \cdots + e_l\beta_l = rs\beta &= s\Lambda_0 - (s\Lambda_0 - rs\beta) \\ &= (m_1 - d_1)\beta_1 + \cdots + (m_l - d_l)\beta_l \end{aligned}$$

and therefore $m_i = d_i - re_i \geqq 0$. Put $d = \max_i d_i$. Then $d \geqq \max_i re_i \geqq r$. This proves that $\pi(X_{-\beta}^{d+1})\psi_0 = 0$ and so $\beta \varepsilon Q'$.

Notice that if α is a positive compact root $\alpha \varepsilon Q'$. For the dimension of $\pi(\mathfrak{X})\psi_0$ is finite and the nonzero vectors among $\pi(X_{-\alpha}^r)\psi_0$ $(r \geqq 0)$ are linearly independent since they belong to the distinct weights $\Lambda_0 - r\alpha$. Hence $\pi(X_{-\alpha}^r)\psi_0 = 0$ for r sufficiently large. Therefore if W_k is the subgroup of W generated by the Weyl reflexions corresponding to the compact roots, $W_k \subset W'$.

Let $\mathfrak{c}_0$ be the center of $\mathfrak{k}_0$ and let $\mathfrak{k}_0' = [\mathfrak{k}_0, \mathfrak{k}_0]$. Then $\mathfrak{k}_0$ is the direct sum of $\mathfrak{c}_0$ and $\mathfrak{k}_0'$ (see [5(c), § 2]) and since $\mathfrak{h}_0$ is maximal abelian in $\mathfrak{k}_0$, $\mathfrak{c}_0 \subset \mathfrak{h}_0$. Let $\mathfrak{c}$ and $\mathfrak{k}'$ be the respective complexifications of $\mathfrak{c}_0$ and $\mathfrak{k}_0'$ in $\mathfrak{k}$. Put $\mathfrak{h}' = \mathfrak{h} \cap \mathfrak{k}'$.

LEMMA 8. *Let β be a positive root. If β is zero on $\mathfrak{c}$, $s\beta < 0$ for some $s \varepsilon W_k$ and therefore $\beta \varepsilon Q'$.*

It is obvious that every compact root is identically zero on $\mathfrak{c}$. Since $\beta - s\beta$ $(s \varepsilon W_k)$ is a linear combination of compact roots it follows that if β vanishes on $\mathfrak{c}$ the same holds for $s\beta$. Now consider $\lambda = \sum_{s \varepsilon W_k} s\beta$. Then $s_\alpha \lambda = \lambda$ and therefore $\lambda(H_\alpha) = 0$ for every compact root α. Since $\mathfrak{h}'$ is a Cartan subalgebra of the semisimple algebra $\mathfrak{k}'$ (see [5(c), Lemma 3]) it follows that the elements H_α corresponding to all compact roots α span $\mathfrak{h}'$. Hence $\lambda = 0$ on $\mathfrak{h}'$. But as we have seen $s\beta$ $(s \varepsilon W_k)$ and therefore also λ vanish on $\mathfrak{c}$. Hence $\lambda = 0$ and therefore $s\beta$ cannot all be positive. This proves that $s\beta < 0$ for some $s \varepsilon W_k$. But $W_k \subset W'$ and so we can conclude from Lemma 7 that $\beta \varepsilon Q'$.

We define the representation σ of $\mathfrak{k}$ on $\mathfrak{g}$ as before.

LEMMA 9. *Let γ be a root of $\mathfrak{g}$. Then $\sigma(\mathfrak{X})X_\gamma$ is irreducible under $\mathfrak{k}$.*

Since $\mathfrak{k}$ is reductive in $\mathfrak{g}$ [5(c), Lemma 3] the representation σ is semisimple. Also it is obvious that $\mathfrak{g}$ is fully reducible under $\sigma(\mathfrak{h})$ and CX_γ is the space of all elements $Y \varepsilon \mathfrak{g}$ such that $\sigma(H)Y = \gamma(H)Y$ for every $H \varepsilon \mathfrak{h}$. Let $\mathfrak{q}$ be a subspace of $\sigma(\mathfrak{X})X_\gamma$ which is invariant under $\sigma(\mathfrak{k})$. Since σ is semisimple there exists a subspace $\mathfrak{q}'$ complementary to $\mathfrak{q}$ in $\sigma(\mathfrak{X})X_\gamma$ which is also invariant under $\sigma(\mathfrak{k})$. Since $X_\gamma \varepsilon \mathfrak{q} + \mathfrak{q}'$, we can choose an element $Y \neq 0$ either in $\mathfrak{q}$ or $\mathfrak{q}'$ such that $\sigma(H)Y = \gamma(H)Y$ for all $H \varepsilon \mathfrak{h}$. Let $\mathfrak{q}_1 = \mathfrak{q}$ or $\mathfrak{q}'$ according as Y is in $\mathfrak{q}$ or $\mathfrak{q}'$. Then as we have remarked above $CY = CX_\gamma$ and therefore $\sigma(\mathfrak{X})X_\gamma = \sigma(\mathfrak{X})Y \subset \mathfrak{q}_1$. This proves that $\mathfrak{q}_1 = \sigma(\mathfrak{X})X_\gamma$ and so $\sigma(\mathfrak{X})X_\gamma$ is irreducible.

COROLLARY. *If* β *is any root such that* $X_\beta \in \sigma(\mathfrak{X})X_\gamma$ *then* $\beta - \gamma$ *is identically zero on* $\mathfrak{c}$.

For if $H \in \mathfrak{c}$, $\sigma(H)$ lies in the center of $\sigma(\mathfrak{X})$. Hence in view of the irreducibility proved above, it follows from Schur's lemma that there exists a complex number c such that $\sigma(H)Y = cY$ for all $Y \in \sigma(\mathfrak{X})X_\gamma$. Hence in particular if $Y = X_\beta$, $cX_\beta = \sigma(H)X_\beta = \beta(H)X_\beta$. Therefore $c = \beta(H)$ and similarly $c = \gamma(H)$. Hence $\beta(H) - \gamma(H) = 0$ for every $H \in \mathfrak{c}$.

Now we come to the proof of Theorem 1. Let γ_0 be a positive root which is not in Q'. We have to show that γ_0 is totally positive. Since $P_k \subset Q'$, γ_0 is not compact and therefore $X_{\gamma_0} \in \mathfrak{p}$. Put $\mathfrak{p}_{\gamma_0} = \sigma(\mathfrak{X})X_{\gamma_0}$. Then $\mathfrak{p}_{\gamma_0} \subset \mathfrak{p}$ and since $\mathfrak{p}$ is fully reducible under $\sigma(\mathfrak{h})$, $\mathfrak{p}_{\gamma_0}$ can be spanned by the elements X_β corresponding to a certain number of noncompact roots β. Let β_0 and β_1 respectively be the highest and lowest among all those roots β for which $X_\beta \in \mathfrak{p}_{\gamma_0}$. Then $\beta_0 \geqq \gamma_0 > 0$ and in order to prove that γ_0 is totally positive it would be enough to show that $\beta_1 > 0$. So suppose this is false and β_1 is negative. Since $\mathfrak{k} = \mathfrak{c} + \mathfrak{k}'$, it is obvious from Schur's lemma and Lemma 9 that $\mathfrak{p}_{\gamma_0}$ is irreducible under $\sigma(\mathfrak{k}')$. Let τ denote the irreducible representation of $\mathfrak{k}'$ on $\mathfrak{p}_{\gamma_0}$ defined by σ. Now $\mathfrak{h}'$ is a Cartan subalgebra of the semisimple algebra $\mathfrak{k}'$ and

$$\mathfrak{k}' = \mathfrak{h}' + \sum_{\alpha \in P_k} (CX_\alpha + CX_{-\alpha}).$$

Hence if we identify linear functions on $\mathfrak{h}'$ with those linear functions on $\mathfrak{h}$ which vanish on $\mathfrak{c}$, the roots of $\mathfrak{k}'$ (with respect to $\mathfrak{h}'$) coincide with the compact roots of $\mathfrak{g}$. For any root β of $\mathfrak{g}$ let β' denote the restriction of β on $\mathfrak{h}'$. It is clear that $[X_\alpha, X_{\beta_0}] = [X_{-\alpha}, X_{\beta_1}] = 0$ for all positive compact roots α. Therefore it follows from Lemma 2 that β_0' is the highest and β_1' the lowest weight (with respect to $\mathfrak{h}'$) of the irreducible representation τ of $\mathfrak{k}'$. Hence from Lemma 3, there exists an element t in the Weyl group of $\mathfrak{k}'$ with respect to $\mathfrak{h}'$, such that $\beta_1' = t\beta_0'$. This means that we can choose $s \in W_k$ such that $\beta_1 - s\beta_0$ is identically zero on $\mathfrak{h}'$. However $\beta_0 - s\beta_0$ is a linear combination of compact roots and therefore it vanishes on $\mathfrak{c}$. On the other hand from the above corollary $\beta_1 - \beta_0$ is also zero on $\mathfrak{c}$. Therefore $\beta_1 - s\beta_0$ vanishes on both $\mathfrak{h}'$ and $\mathfrak{c}$ and so $\beta_1 = s\beta_0$. Since $\beta_0 > 0$ and $s\beta_0 = \beta_1 < 0$, we conclude from Lemma 7 that $\beta_0 \in Q'$.

Now $X_{\gamma_0} \in \mathfrak{p}_{\gamma_0}$ and $\gamma_0 \notin Q'$. Choose the highest β such that $X_\beta \in \mathfrak{p}_{\gamma_0}$ and $\beta \notin Q'$. Then $\beta \geqq \gamma_0 > 0$ and $\beta \neq \beta_0$. We claim that $[X_\alpha, X_\beta] \neq 0$ for some positive compact root α. For otherwise suppose $[X_\alpha, X_\beta] = 0$ for every $\alpha \in P_k$. Let $\alpha_1, \cdots, \alpha_r$ be all the distinct positive compact roots of $\mathfrak{g}$. Since $\mathfrak{p}_{\gamma_0}$ is irreducible under $\sigma(\mathfrak{k})$, it follows that $\sigma(\mathfrak{X})X_\beta = \mathfrak{p}_{\gamma_0}$. But then from Lemma 2 and the corollary to Lemma 9 we can conclude that

$$\beta_0 = \beta - (m_1\alpha_1 + \cdots + m_r\alpha_r) \qquad (m_i \geqq 0).$$

Since $\beta_0 \geqq \beta$ this implies that $m_1 = \cdots = m_r = 0$. However this is impossible as $\beta \neq \beta_0$. So let α be a positive compact root such that $[X_\alpha, X_\beta] \neq 0$. Then $\beta + \alpha$ is a root and $X_{\beta+\alpha} = c[X_\alpha, X_\beta] \varepsilon \mathfrak{p}_{\gamma_0}$ $(c \varepsilon C)$. Moreover $\beta + \alpha > \beta$ and therefore $\beta \varepsilon Q'$. Now $\beta + (\beta + \alpha)$ is not a root. For since $\beta + \alpha$ and β are both noncompact $\beta + (\beta + \alpha)$, if it is a root, must be compact. On the other hand $\beta - (\beta + \alpha) = -\alpha$ is also a compact root. Since $2\beta = \beta + (\beta + \alpha) - \alpha$, β is a linear combination of two compact roots and so it must be zero on $\mathfrak{c}$. Therefore from Lemma 8, $\beta \varepsilon Q'$. As this contradicts the definition of β, it follows that $\beta + (\beta + \alpha)$ is not a root while $\beta - (\beta + \alpha)$ is a root. Let ν be the largest integer $\geqq 0$ such that $\beta - \nu(\beta + \alpha)$ is a root. Then $\nu \geqq 1$ and

$$s_{\beta+\alpha}\beta = \beta - \nu(\beta + \alpha) \leqq \beta - (\beta + \alpha) = -\alpha < 0.$$

But $\beta + \alpha \varepsilon Q'$ and therefore from Lemma 7, $\beta \varepsilon Q'$ giving again a contradiction. This proves that $\beta_1 > 0$ and so γ_0 is totally positive.

4. Another characterization of totally positive roots.

We shall now study the totally positive roots a little more closely. We call a root β totally negative if $-\beta$ is totally positive.

LEMMA 10. *Let β be a totally positive root. Then $s\beta$ $(s \varepsilon W_k)$ is also totally positive. Similarly if β is totally negative the same is true of $s\beta$.*

Let α be a compact root and let $\gamma = s_\alpha\beta$. Then it is known (see Weyl [7]) that X_γ can be written in the form $X_\gamma = c\sigma(Y^m)X_\beta$ where $c \varepsilon C$ and $Y = X_\alpha$ or $X_{-\alpha}$ and m is a nonnegative integer. Therefore from Lemma 9, $\sigma(\mathfrak{X})X_\gamma = \sigma(\mathfrak{X})X_\beta$. This shows that if β is totally positive the same holds for $s_\alpha\beta$ and therefore also for $s\beta$ $(s \varepsilon W_k)$. On the other hand if β is totally negative $s(-\beta) = -s\beta$ $(s \varepsilon W_k)$ is totally positive and therefore $s\beta$ is totally negative.

COROLLARY 1. *If β is totally positive it cannot be identically zero on* $\mathfrak{c}$.

Put $\lambda = \sum_{s \varepsilon W_k} s\beta$. Then $s\beta$ are all totally positive and hence positive. Therefore $\lambda > 0$. But $s_\alpha\lambda = \lambda$ for every compact root α and so λ is zero on $\mathfrak{h}'$. Hence λ cannot be identically zero on $\mathfrak{c}$. But $\beta - s\beta$ $(s \varepsilon W_k)$ being a linear combination of compact roots is zero on $\mathfrak{c}$. This shows that if w

is the order of the group W_k, λ coincides with $w\beta$ on $\mathfrak{c}$. Therefore β cannot vanish identically on $\mathfrak{c}$.

COROLLARY 2. *Every totally positive root is noncompact. If* $\mathfrak{c} = \{0\}$ *there are no totally positive roots.*

This is an immediate consequence of Corollary 1 above.

LEMMA 11. *Let* β *be a totally positive root and* γ *a noncompact root which is not totally negative. Then* $\beta + \gamma$ *is not a root.*

For otherwise suppose $\beta + \gamma$ is a root. Since β and γ are both noncompact $\alpha = \beta + \gamma$ must be compact. Then $-\gamma = \beta - \alpha$ and

$$X_{-\gamma} = c[X_{-\alpha}, X_\beta] \varepsilon \sigma(\mathfrak{X}) X_\beta \qquad (c \varepsilon C).$$

Hence from Lemma 9, $\sigma(\mathfrak{X}) X_{-\gamma} = \sigma(\mathfrak{X}) X_\beta$ and therefore since β is totally positive the same must hold for $-\gamma$. But this means that γ is totally negative and so we get a contradiction.

Let θ denote the automorphism of $\mathfrak{g}$ of order 2 given by $\theta(X + Y) = X - Y$ $(X \varepsilon \mathfrak{k}, Y \varepsilon \mathfrak{p})$. We define a linear mapping $\bar{\theta}$ of $\mathfrak{g}$ over R by the rule $\bar{\theta}(cX) = \bar{c}\theta(X)$ $(c \varepsilon C, X \varepsilon \mathfrak{g}_0)$ where $\bar{c}$ is the complex conjugate of c. Then $\bar{\theta}$ is the conjugation of $\mathfrak{g}$ with respect to its compact real form[5] $\mathfrak{u} = \mathfrak{k}_0 + (-1)^{\frac{1}{2}}\mathfrak{p}_0$ (see [5(c), § 2]) and it is known (see Weyl [7]) that one can choose the elements X_γ corresponding to the various roots γ in such a way that $X_\gamma - X_{-\gamma}$ and $(-1)^{\frac{1}{2}}(X_\gamma + X_{-\gamma})$ both lie in $\mathfrak{u}$ (without disturbing our earlier condition that $\gamma(H_\gamma) = 2$ for $H_\gamma = [X_\gamma, X_{-\gamma}]$). Assuming that this has been done, it follows immediately that $\bar{\theta}(X_\gamma) = -\epsilon_\gamma X_{-\gamma}$ where $\epsilon_\gamma = 1$ or -1 according as γ is compact or not.

LEMMA 12. *Let* $\alpha_1, \cdots, \alpha_r$ *be all the distinct positive compact roots and* β *a totally positive root. Then if* γ *is any root of the form* $\beta + m_1\alpha_1 + \cdots + m_r\alpha_r$ *(where* $m_1, \cdots, m_r$ *are integers),* γ *is totally positive and* $X_\gamma \varepsilon \sigma(\mathfrak{X}) X_\beta$.

We shall make use of an argument due to E. Cartan [3(c)]. Define the bilinear form $B(X, Y)$ on $\mathfrak{g}$ as before. Then $-B(X, \bar{\theta}(X))$ $(X \varepsilon \mathfrak{g})$ is a positive definite Hermitian form on $\mathfrak{g}$. Since β is noncompact, $\mathfrak{p}_\beta = \sigma(\mathfrak{X}) X_\beta$ is a subspace of $\mathfrak{p}$. Let $\mathfrak{q}$ denote the orthogonal complement[6] of $\mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta)$ in $\mathfrak{p}$ (with respect to the above Hermitian form). Then we shall prove that $[\mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta), \mathfrak{q}] = \{0\}$. Since $\mathfrak{q} = \tilde{\theta}(\mathfrak{q})$, it is enough to show that $[\mathfrak{p}_\beta, \mathfrak{q}] = \{0\}$. Let $X \varepsilon \mathfrak{p}_\beta$, $Y \varepsilon \mathfrak{q}$ and $Z \varepsilon \mathfrak{k}$. Then

---

[5] We fix once for all one square-root of -1 in C and denote it by $(-1)^{\frac{1}{2}}$.

[6] I take this opportunity to correct a slip in one of my earlier papers [5(d), p. 29 bottom paragraph]. There the orthogonal complement is taken with respect to $B(X, Y)$ instead of $B(X, \bar{\theta}, Y)$). This is a mistake which however is easy to rectify without making any substantial change in the argument.

$$B([X,Y], \tilde{\theta}(Z)) = B(Y, [\tilde{\theta}(Z), X]) = 0$$

since $[\theta(\tilde{Z}), X]$ lies in $\mathfrak{p}_\beta$ and $\mathfrak{q}$ is orthogonal to $\tilde{\theta}(\mathfrak{p}_\beta)$. But $[X,Y] \varepsilon [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}$ and so this shows that $[X,Y] = 0$ and our assertion is proved. Let $\mathfrak{k}_\beta$ be the centralizer of $\mathfrak{q}$ in $\mathfrak{k}$. Put $\mathfrak{g}_\beta = \mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta) + \mathfrak{k}_\beta$. We claim that $\mathfrak{g}_\beta$ is an ideal in $\mathfrak{g}$. Since $\mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta)$ is invariant under $\sigma(\mathfrak{k})$, the same holds for $\mathfrak{q}$ and $\mathfrak{k}_\beta$. Therefore $[\mathfrak{k}, \mathfrak{g}_\beta] \subset \mathfrak{g}_\beta$. Moreover $[\mathfrak{q}, \mathfrak{g}_\beta] = \{0\}$. Since $\mathfrak{g} = \mathfrak{k} + \mathfrak{q} + \mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta)$, it is enough to prove that $[\mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta), \mathfrak{g}_\beta] \subset \mathfrak{g}_\beta$. But $\mathfrak{g}_\beta = \mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta) + \mathfrak{k}_\beta$ and $\mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta)$ is invariant under $\sigma(\mathfrak{k})$. So it would be sufficient to show that if X, Y are in $\mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta)$, $[X,Y]$ lies in $\mathfrak{k}_\beta$. However $[X,Y] \varepsilon \mathfrak{k}$ and if $Z \varepsilon \mathfrak{q}$

$$[[X,Y],Z] = [[X,Z],Y] + [X,[Y,Z]] = 0$$

since $[X,Z] = [Y,Z] = 0$. This proves that $[X,Y] \varepsilon \mathfrak{k}_\beta$ and therefore $\mathfrak{g}_\beta$ is an ideal in $\mathfrak{g}$. Let $\mathfrak{l}$ be the orthogonal complement of $\mathfrak{k}_\beta$ in $\mathfrak{k}$. Then $\mathfrak{g}' = \mathfrak{q} + \mathfrak{l}$ is clearly the orthogonal complement of $\mathfrak{g}_\beta$ in $\mathfrak{g}$. Let $X \varepsilon \mathfrak{g}$ and $Y \varepsilon \mathfrak{g}'$. Then if $Z \varepsilon \mathfrak{g}_\beta$,

$$B([X,Y], \tilde{\theta}(Z)) = -B(Y, [X, \tilde{\theta}(Z)]).$$

But

$$\tilde{\theta}([X, \tilde{\theta}(Z)]) = [\tilde{\theta}(X), Z] \varepsilon \mathfrak{g}_\beta$$

and since $\mathfrak{g}'$ is orthogonal to $\mathfrak{g}_\beta$, $B(Y, [X, \tilde{\theta}(Z)]) = 0$. Hence $[X,Y]$ being orthogonal to $\mathfrak{g}_\beta$ lies in $\mathfrak{g}'$. This shows that $\mathfrak{g}'$ is also an ideal in $\mathfrak{g}$. Hence $[\mathfrak{g}', \mathfrak{g}_\beta] \subset \mathfrak{g}' \cap \mathfrak{g}_\beta = \{0\}$. Moreover since $\mathfrak{q} = \tilde{\theta}(\mathfrak{q})$ it follows that $\mathfrak{k}_\beta = \tilde{\theta}(\mathfrak{k}_\beta)$, $\tilde{\theta}(\mathfrak{g}_\beta) = \mathfrak{g}_\beta$ and $\tilde{\theta}(\mathfrak{g}') = \mathfrak{g}'$. Now suppose $X \varepsilon \mathfrak{c}$. Since $\mathfrak{k} = \mathfrak{k}_\beta + \mathfrak{l}$, $X = Y + Z$ where $Y \varepsilon \mathfrak{k}_\beta$ and $Z \varepsilon \mathfrak{l}$. But $[\mathfrak{k}_\beta, \mathfrak{l}] = \{0\}$ and so it is obvious that both Y and Z are in $\mathfrak{c}$. This shows that $\mathfrak{c} = \mathfrak{c} \cap \mathfrak{k}_\beta + \mathfrak{c} \cap \mathfrak{l}$.

Now suppose $\gamma = \beta + m_1\alpha_1 + \cdot\cdot\cdot + m_r\alpha_r$ is a root. Then γ and β coincide on $\mathfrak{c}$. Since $X_\beta \varepsilon \mathfrak{g}_\beta$, it commutes with $\mathfrak{l}$ and therefore β vanishes identically on $\mathfrak{c} \cap \mathfrak{l}$. Hence the same holds for γ. Moreover since both $\mathfrak{g}_\beta$ and $\mathfrak{g}'$ are invariant under $\sigma(\mathfrak{h})$ it is clear that X_γ lies either in $\mathfrak{g}_\beta$ or $\mathfrak{g}'$. We claim $X_\gamma \varepsilon \mathfrak{g}_\beta$. For otherwise if $X_\gamma \varepsilon \mathfrak{g}'$ it commutes with $\mathfrak{g}_\beta$ and therefore γ vanishes identically on $\mathfrak{c} \cap \mathfrak{k}_\beta$. But we have already seen that γ is zero on $\mathfrak{c} \cap \mathfrak{l}$. Therefore γ and hence also β vanishes identically on $\mathfrak{c}$. But since β is totally positive this is impossible in view of Corollary 1 to Lemma 10. This proves that $X_\gamma \varepsilon \mathfrak{g}_\beta$. Now γ cannot be a compact root since it coincides with β on $\mathfrak{c}$ and therefore is not identically zero on $\mathfrak{c}$. Hence $X_\gamma \varepsilon\, \mathfrak{p} \cap \mathfrak{g}_\beta = \mathfrak{p}_\beta + \tilde{\theta}(\mathfrak{p}_\beta)$. Since $\mathfrak{p}_\beta$ and $\tilde{\theta}(\mathfrak{p}_\beta)$ are both invariant under $\sigma(\mathfrak{h})$ either $X_\gamma \varepsilon\, \mathfrak{p}_\beta$ or $X_\gamma \varepsilon\, \tilde{\theta}(\mathfrak{p}_\beta)$. But $\tilde{\theta}(X_\gamma) = -\epsilon_\gamma X_{-\gamma}$ and so either X_γ or $X_{-\gamma}$ lies in $\mathfrak{p}_\beta$. Now if $X_{-\gamma} \varepsilon\, \mathfrak{p}_\beta$, it follows from the corollary to Lemma 9 that $\beta + \gamma$ is zero on $\mathfrak{c}$. But we have seen above that $\beta - \gamma$ is zero on $\mathfrak{c}$. Therefore $2\beta = (\beta + \gamma) + (\beta - \gamma)$ is also zero on $\mathfrak{c}$. Since β is totally positive this is impossible (Corollary 1 to Lemma 10) and therefore $X_\gamma \varepsilon\, \mathfrak{p}_\beta$. Hence from Lemma 9 $\sigma(\mathfrak{X})X_\gamma = \sigma(\mathfrak{X})X_\beta$ and this proves that γ is totally positive.

COROLLARY. *Let* $\alpha_1, \cdots, \alpha_r$ *be all the positive compact roots of* $\mathfrak{g}$. *Then a root* β *is totally positive if and only if every root of the form* $\beta + m_1\alpha_1 + \cdots + m_r\alpha_r$ (where $m_1, \cdots, m_r$ are integers) *is positive.*

Suppose $\gamma = \beta + m_1\alpha_1 + \cdots + m_r\alpha_r$ is a root. If β is totally positive, it follows from the above lemma that γ is positive. Conversely suppose every root γ of this form is positive. Then it is clear that $\beta \neq \pm\alpha_i$, $1 \leqq i \leqq r$ and therefore β is noncompact and $\sigma(\mathfrak{X})X_\beta$ is spanned by the elements X_γ corresponding to certain roots γ of the above form. Since every such γ is positive, it follows from our definition that β is totally positive. The above corollary thus gives an alternative definition of total positivity.

5. Some further results on totally positive roots.

First we recall a few well-known facts in the theory of semisimple Lie algebras. Let $\mathfrak{a}$ be an ideal in $\mathfrak{g}$ and β a root of $\mathfrak{g}$. We say that β is a root of $\mathfrak{a}$ if $X_\beta \varepsilon\, \mathfrak{a}$. Let $(\beta_1, \cdots, \beta_l)$ be a fundamental system of roots of $\mathfrak{g}$ and suppose among these $\beta_1, \cdots, \beta_m$ are all the roots of $\mathfrak{a}$. Then every root of $\mathfrak{a}$ is a linear combination of $\beta_1, \cdots, \beta_m$. Let s_i denote the Weyl reflexion corresponding to β_i and let $W_\mathfrak{a}$ be the subgroup of the Weyl group W generated by $s_1, \cdots, s_m$. Then if β is any root of $\mathfrak{a}$ the corresponding Weyl reflexion s_β is contained in $W_\mathfrak{a}$.

Now as before let P_+ be the set of all totally positive roots. Put $\mathfrak{p}_+ = \sum_{\beta \varepsilon P_+} CX_\beta$ and $\mathfrak{p}_- = \sum_{\beta \varepsilon P_+} CX_{-\beta}$ so that $\mathfrak{p}_- = \tilde{\theta}(\mathfrak{p}_+)$. Let $\mathfrak{q}$ be the orthogonal complement of $\mathfrak{p}_+ + \mathfrak{p}_-$ in $\mathfrak{p}$ (under the Hermitian form $B(X, \tilde{\theta}(X))$). Then if $\mathfrak{l}$ is the centralizer of $\mathfrak{q}$ in $\mathfrak{k}$ we can show, as in the proof of Lemma 12, that $\mathfrak{g}_+ = \mathfrak{p}_+ + \mathfrak{p}_- + \mathfrak{l}$ is an ideal in $\mathfrak{g}$. Let $\mathfrak{g}'$ be orthogonal complement of $\mathfrak{g}_+$ in $\mathfrak{g}$. Then $\mathfrak{g}'$ is also an ideal in $\mathfrak{g}$ and $[\mathfrak{g}', \mathfrak{g}_+] = \{0\}$. It is obvious that $\mathfrak{g}' \cap \mathfrak{k}$ and $\mathfrak{g}' \cap \mathfrak{p}$ are the orthogonal complements of $\mathfrak{l}$ and $\mathfrak{p}_+ + \mathfrak{p}_-$ in $\mathfrak{k}$ and $\mathfrak{p}$ respectively and $\mathfrak{g}' = \mathfrak{g}' \cap \mathfrak{k} + \mathfrak{g}' \cap \mathfrak{p}$. Now if β is a root, X_β lies either in $\mathfrak{g}_+$ or $\mathfrak{g}'$. Since $\mathfrak{g}_+ \cap \mathfrak{p} = \mathfrak{p}_+ + \mathfrak{p}_-$, it follows that if β is a noncompact root of $\mathfrak{g}_+$ it is either totally positive or totally negative.

Let $(\beta_1, \cdots, \beta_l)$ be a fundamental system of positive roots of $\mathfrak{g}$ (Corollary 2 to Lemma 4). If γ is a positive root which is not contained in this fundamental system, then it is well known that $\gamma - \beta_i$ must also be a positive root for some i $(1 \leqq i \leqq l)$. On the other hand $\beta_i - \beta$ can never be a positive root for any root $\beta > 0$. Hence a positive root is contained in the set $(\beta_1, \cdots, \beta_l)$ if and only if $\gamma - \beta$ is never a positive root for any root $\beta > 0$.

Now it follows from the definition of total positivity that $\mathfrak{p}_+$ is invariant under $\sigma(\mathfrak{k})$. Hence $\mathfrak{p}_+ = \sum_{\nu=1}^{q} \mathfrak{q}_\nu$ where $\mathfrak{q}_\nu$, $1 \leqq \nu \leqq q$, are nonzero subspaces of $\mathfrak{p}_+$ which are invariant and irreducible under $\sigma(\mathfrak{k})$ and the sum is direct. (If $\mathfrak{p}_+ = \{0\}$, $q = 0$). Let γ_ν be the lowest root such that $X_{\gamma_\nu} \varepsilon \mathfrak{q}_\nu$ and let α be a positive compact root. We claim $\gamma_\nu - \alpha$ is not a root. For otherwise it follows from Lemma 12 that $X_{\gamma_\nu - \alpha} \varepsilon \mathfrak{q}_\nu$ which however is impossible since $\gamma_\nu - \alpha < \gamma_\nu$. Now let β be a noncompact positive root and suppose $\alpha = \gamma_\nu - \beta$ is a root. Since γ_ν and β are both noncompact, α is compact and $\beta = \gamma_\nu - \alpha$. Hence by the above argument α cannot be positive. This proves that $\gamma_\nu - \beta$ can never be a positive root for any root $\beta > 0$. Therefore it follows from our criterion above that γ_ν is a root of the fundamental system $(\beta_1, \cdots, \beta_l)$.

Now suppose $(\beta_1, \cdots, \beta_r)$ are roots of $\mathfrak{g}_+$ while $(\beta_{r+1}, \cdots, \beta_l)$ are roots of $\mathfrak{g}'$ $(0 \leqq r \leqq l)$. Then if α is a positive compact root of $\mathfrak{g}_+$,

$$\alpha = m_1\beta_1 + \cdots + m_r\beta_r \qquad (m_i \geqq 0).$$

We recall that if γ is any compact root $\sum_{s \varepsilon W_k} s\gamma = 0$. Now suppose $\beta_1, \cdots, \beta_t$ are compact while $\beta_{t+1}, \cdots, \beta_r$ are noncompact (and hence totally positive). Then

$$0 = \sum_{s \varepsilon W_k} s\alpha = \sum_{i=1}^{r} m_i \sum_{s \varepsilon W_k} s\beta_i = \sum_{t < i \leqq r} m_i \sum_{s \varepsilon W_k} s\beta_i$$

But from Lemma 10 $s\beta_i$ $(i > t)$ is also totally positive. Since $m_i \geqq 0$ the above sum cannot be zero unless $m_i = 0$ $(t < i \leqq r)$. Therefore $\alpha = m_1\beta_1 + \cdots + m_t\beta_t$. On the other hand let γ be any positive noncompact root of $\mathfrak{g}_+$. Then $X_\gamma \varepsilon \mathfrak{q}_\nu = \sigma(\mathfrak{X}) X_{\gamma_\nu}$ for some ν. Since $[\mathfrak{g}', \mathfrak{g}_+] = \{0\}$, it is clear that $\mathfrak{q}_\nu$ is irreducible under $\sigma(\mathfrak{l})$ and therefore $\gamma - \gamma_\nu$ is a linear combination of compact roots of $\mathfrak{g}_+$. Hence $\gamma = \gamma_\nu + m_1\beta_1 + \cdots + m_t\beta_t$ where $m_1, \cdots, m_t$ are real numbers. But γ_ν is contained in the set $(\beta_1, \cdots, \beta_r)$ and since it is noncompact, $\gamma_\nu \neq \beta_i$, $1 \leqq i \leqq t$. Therefore since $\gamma > 0$ and $(\beta_1, \cdots, \beta_l)$ is a fundamental system, $m_1, \cdots, m_l$ must all be nonnegative integers. Thus we have obtained the following result.

LEMMA 13. *Let* $(\beta_1, \cdots, \beta_l)$ *be a fundamental system of positive roots of* $\mathfrak{g}$. *Let* $(\alpha_1, \cdots, \alpha_r)$ *be all the (distinct) compact roots of* $\mathfrak{g}_+$ *and* $(\gamma_1, \cdots, \gamma_t)$ *all the (distinct) noncompact roots of* $\mathfrak{g}_+$ *which occur in this fundamental system. Then every positive root* γ *of* $\mathfrak{g}_+$ *can be written in the form*

$$\gamma = \mu_1\gamma_1 + \cdots + \mu_t\gamma_t + m_1\alpha_1 + \cdots + m_r\alpha_r$$

where μ_i, m_j *are all nonnegative integers. Moreover* $\mu_1 + \cdots + \mu_t = 0$ *or* 1 *according as* γ *is compact or not.*

COROLLARY 1. $t = \dim(\mathfrak{c} \cap \mathfrak{g}_+)$ *and* $r = \dim(\mathfrak{h}' \cap \mathfrak{g}_+)$.

For it is easy to see that $\mathfrak{h} = \mathfrak{h} \cap \mathfrak{g}_+ + \mathfrak{h} \cap \mathfrak{g}'$ and $\mathfrak{c} = \mathfrak{c} \cap \mathfrak{g}_+ + \mathfrak{c} \cap \mathfrak{g}'$. From this it follows that $r + t = \dim(\mathfrak{h} \cap \mathfrak{g}_+)$ and $r = \dim(\mathfrak{h}' \cap \mathfrak{g}_+)$. Since $\mathfrak{h} \cap \mathfrak{g}_+ = \mathfrak{h}' \cap \mathfrak{g}_+ + \mathfrak{c} \cap \mathfrak{g}_+$, $t = \dim(\mathfrak{c} \cap \mathfrak{g}_+)$.

COROLLARY 2. *If the algebra* $\mathfrak{g}_0$ *is simple, only the following two cases are possible.*

(1) *There are no totally positive roots.*

(2) *Every noncompact positive root is totally positive.*

Suppose we are in case (2) *and* $\mathfrak{g}_0 \neq \mathfrak{k}_0$. *Then there is exactly one noncompact root in the fundamental system* $(\beta_1, \cdots, \beta_l)$ *and* $\dim_R \mathfrak{c}_0 = 1$.

Let η denote the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$ so that

$$\eta(X + (-1)^{\frac{1}{2}}Y) = X - (-1)^{\frac{1}{2}}Y \qquad (X, Y \varepsilon \mathfrak{g}_0).$$

Then $\eta = \tilde{\theta} \circ \theta$ and since $\mathfrak{g}_+$ and $\mathfrak{g}'$ are invariant under both $\tilde{\theta}$ and θ they are also invariant under η. Therefore $\mathfrak{g}_0 = \mathfrak{g}_0 \cap \mathfrak{g}_+ + \mathfrak{g}_0 \cap \mathfrak{g}'$ and $\mathfrak{g}_+$, $\mathfrak{g}'$ are spanned over C by $\mathfrak{g}_0 \cap \mathfrak{g}_+$ and $\mathfrak{g}_0 \cap \mathfrak{g}'$ respectively. Hence it follows that if $\mathfrak{g}_0$ is simple either $\mathfrak{g}_+$ or $\mathfrak{g}'$ is $\{0\}$ and so we get the above two cases. Now suppose we are in case (2). Then every positive root is either compact or totally positive. Therefore if $\mathfrak{g}_0 \neq \mathfrak{k}_0$, totally positive roots must exist and so $t \geqq 1$ in the notation of Lemma 13. On the other hand it follows from the simplicity of $\mathfrak{g}_0$ (see Cartan [3(e), p. 247]) that $\dim_R \mathfrak{c}_0 \leqq 1$. Therefore from Corollary 1 above $t = \dim_R \mathfrak{c}_0 = 1$.

LEMMA 14. *Let* $(\beta_1, \cdots, \beta_q)$ *be all those roots in the fundamental system* $(\beta_1, \cdots, \beta_l)$ *(of Lemma* 13) *which are not totally positive and let* W' *be the subgroup of the Weyl group* W *generated by the Weyl reflexions* *corresponding to* β_i, $1 \leqq i \leqq q$. *Then if* β *is any positive root of* $\mathfrak{g}$ *which is not totally positive we can find an index* i $(1 \leqq i \leqq q)$ *and an element* $s \varepsilon W'$ *such that* $\beta = s\beta_i$.

For any root γ let s_γ denote the corresponding Weyl reflexion. Put $s_i = s_{\beta_i}$, $1 \leqq i \leqq l$. Then it is well known that $s_1, \cdots, s_l$ generate W. Now if some β_i is a root of $\mathfrak{g}'$, it cannot be totally positive and therefore $i \leqq q$. So we may suppose that $(\beta_1, \cdots, \beta_d)$ $(0 \leqq d \leqq q)$ are all the roots of $\mathfrak{g}'$ among $(\beta_1, \cdots, \beta_l)$. Since $[\mathfrak{g}_+, \mathfrak{g}'] = \{0\}$, it is easy to see that if γ and γ' are roots of $\mathfrak{g}_+$ and $\mathfrak{g}'$ respectively then s_γ and $s_{\gamma'}$ commute and $s_\gamma\gamma' = \gamma'$, $s_{\gamma'}\gamma = \gamma$. Therefore in particular s_i and s_j commute if $i \leqq d$ and $j > d$. Let W'' and W_+ be the subgroups of W generated by s_i, $1 \leqq i \leqq d$ and s_j, $d < j \leqq l$ respectively and let β be a positive root which is not totally positive. First suppose β is a root of $\mathfrak{g}'$. Then $\beta = s\beta_i$ for some $s \varepsilon W$ and some i $(1 \leqq i \leqq l)$. But we can write $s^{-1} = uv$ where $u \varepsilon W''$ and $v \varepsilon W_+$. Then $v\beta = \beta$ and therefore $\beta_i = s^{-1}\beta = u\beta$. Since $u \varepsilon W''$, $u\beta - \beta$, and therefore also β_i, is a linear combination of roots of $\mathfrak{g}'$. This proves that β_i is a root of $\mathfrak{g}'$ and therefore $i \leqq d$. Since $u^{-1} \varepsilon W'' \subset W'$ and $\beta = u^{-1}\beta_i$, our assertion is true in this case. On the other hand if β is a root of $\mathfrak{g}_+$, it must be compact. Now $\mathfrak{k}' \cap \mathfrak{g}_+$ being an ideal in the semisimple algebra $\mathfrak{k}'$, is semisimple and the compact roots of $\mathfrak{g}_+$ may be identified in the obvious manner with the roots of $\mathfrak{k}' \cap \mathfrak{g}_+$ with respect to $\mathfrak{h}' \cap \mathfrak{g}_+$. Then it follows from Lemma 13 that $(\beta_{d+1}, \cdots, \beta_q)$ is a fundamental system of roots of $\mathfrak{k}' \cap \mathfrak{g}_+$ and therefore $\beta = s\beta_j$ where $d < j \leqq q$ and s lies in the group generated by $(s_{d+1}, \cdots, s_q)$. This completes the proof of the lemma.

We shall now prove a result which will be needed in a subsequent paper.

LEMMA 15. *Let* γ *be a totally positive root and* α *a compact root. Let* k *and* k' *respectively be the largest nonnegative integers such that* $\gamma - k\alpha$ *and* $\gamma + k'\alpha$ *are roots. Then* $k + k' \leqq 2$.

Put $\gamma_0 = \gamma + k'\alpha$. Then it is known (see Weyl [7]) that $s_\alpha\gamma_0 = \gamma - k\alpha$ and $\gamma_0 - \nu\alpha$ is a root for an integer ν if and only if $0 \leqq \nu \leqq k + k' = \gamma_0(H_\alpha)$. So we have to prove that $\gamma_0(H_\alpha) \leqq 2$. Hence let us suppose that $\gamma_0(H_\alpha) \geqq 3$. Then $\gamma_0 - \alpha$ is a root and therefore also $\alpha - \gamma_0$. Hence α and $\alpha - \gamma_0$ are roots while $\alpha + \gamma_0$ is not a root. This implies (see Weyl [7]) that $s_{\gamma_0}\alpha \leqq \alpha - \gamma_0$ or $\alpha(H_{\gamma_0}) \geqq 1$ and therefore $\gamma_0(H_\alpha)\alpha(H_{\gamma_0}) \geqq 3$. Now consider the root $\gamma' = s_{\gamma_0}s_\alpha\gamma_0$. It is easy to verify that $\gamma' = m\gamma_0 - \gamma_0(H_\alpha)\alpha$ where $m = \gamma_0(H_\alpha)\alpha(H_{\gamma_0}) - 1$. Since γ is totally positive the same holds for γ_0 and $s_\alpha\gamma_0$ (corollary to Lemma 12) and therefore $\gamma' = s_{\gamma_0}s_\alpha\gamma_0$ is a root of $\mathfrak{g}_+$.

However $m \geqq 2$ and so we get a contradiction with Lemma 13. This proves that $\gamma_0(H_\alpha) \leqq 2$.

6. The converse of Theorem 1. Assuming that we are given an order in $\mathfrak{F}_R$ we shall now pove the following converse of Theorem 1.

THEOREM 2. *Let Λ_0 be a linear function on $\mathfrak{h}$ such that $\Lambda_0(H_\beta)$ is a nonnegative integer for every positive root β which is not totally positive. Then there exists a representation π of $\mathfrak{B}$ on a vector space V and a vector $\psi \neq 0$ in V such that the following condition holds:*

(1) *$\pi(X_\beta)\psi = 0$ for every positive root β.*

(2) *$\pi(H)\psi = \Lambda_0(H)\psi$ for all H in $\mathfrak{h}$.*

(3) *π is irreducible and $\dim \pi(\mathfrak{X})\psi$ is finite.*

Moreover such a representation is unique up to equivalence.

Choose a fundamental system $(\beta_1, \cdots, \beta_l)$ of positive roots and let $(\beta_1, \cdots, \beta_q)$ be all the roots in this system which are not totally positive. Then in the notation of Lemma 13, $(\gamma_1, \cdots, \gamma_t)$ are exactly the remaining roots of our fundamental system. Consider the left ideal

$$\mathfrak{M}_0 = \sum_{\beta \varepsilon P} \mathfrak{B} X_\beta + \sum_{\beta \varepsilon P} \mathfrak{B}(H_\beta - \Lambda_0(H_\beta)) + \sum_{1 \leqq i \leqq q} \mathfrak{B}\, Y_i^{\lambda_i + 1}$$

in $\mathfrak{B}$. Here P is the set of all positive roots, $\lambda_i = \Lambda_0(H_{\beta_i})$ and $Y_i = X_{-\beta_i}$, $1 \leqq i \leqq q$. First we claim that $\mathfrak{M}_0 \neq \mathfrak{B}$ (cf. Lemma 8 of [5(b)]). For otherwise if $\mathfrak{M}_0 = \mathfrak{B}$ we can choose $z_i \varepsilon \mathfrak{B}$ $(1 \leqq i \leqq q)$ such that

$$1 \equiv \sum_{1 \leqq i \leqq q} z_i Y_i^{\lambda_i + 1} \bmod \mathfrak{P}_{\Lambda_0}$$

where $\mathfrak{P}_{\Lambda_0} = \sum_{\beta \varepsilon P} \mathfrak{B} X_\beta + \sum_{\beta \varepsilon P} \mathfrak{B}(H_\beta - \Lambda\ (H_\beta))$. Since 1 is of rank zero (see Section 2 for the definition of rank) while $Y_i^{\lambda_i+1}$ is of rank $-(\lambda_i + 1)\beta_i$, we may assume (see the proof of Lemma 8 of [5(b)]) that z_i is of rank $(\lambda_i + 1)\beta_i$. Put $X_i = X_{\beta_i}$, $Y_i = X_{-\beta_i}$ and $H_i = H_{\beta_i}$ $(1 \leqq i \leqq l)$. Then $(1, X_i, H_i, Y_i)_{1 \leqq i \leqq l}$ generate $\mathfrak{B}$ (Lemma 18 of [5(b)]). Therefore if $\mathfrak{U}, \mathfrak{H}, \mathfrak{B}$ are the subalgebras of $\mathfrak{B}$ generated by $(1, X_1, \cdots, X_l)$, $(1, H_1, \cdots, H_l)$ and $(1, Y_1, \cdots, Y_l)$ respectively, it is easy to see that $\mathfrak{B} = \mathfrak{B}\mathfrak{H}\mathfrak{U}$ (see the proof of Lemma 2 of [5(b)]). However in view of the argument given in the proof of Lemma 8 of [5(b), p. 37] we may assume that $z_i \varepsilon \mathfrak{U}$. Since the product $X_{i_1} X_{i_2} \cdots X_{i_r}$ is of rank $\beta_{i_1} + \beta_{i_2} + \cdots \beta_{i_r}$ and since $\beta_1, \cdots, \beta_l$ are linearly independent, it follows that the only elements in $\mathfrak{U}$ of rank $(\lambda_i + 1)\beta_i$ are the numerical multiples of $X_i^{\lambda_i+1}$. Therefore $z_i = c_i X_i^{\lambda_i+1}$ $(c_i \varepsilon C)$. But from Lemma 6 of [5(b)],

$$X_i^{\lambda_i+1}Y_i^{\lambda_i+1} = X_i^{\lambda_i}[X_i, Y_i^{\lambda_i+1}] + X_i^{\lambda_i}Y_i^{\lambda_i+1}X_i$$
$$= \lambda_i X_i^{\lambda_i}Y_i^{\lambda_i}(H_i - \lambda_i) + X_i^{\lambda_i}Y_i^{\lambda_i+1}X_i \equiv 0 \mod \mathfrak{P}_{\Lambda_0} \quad (1 \leqq i \leqq q).$$

Therefore $1 \equiv 0 \mod \mathfrak{P}_{\Lambda_0}$ and so $\mathfrak{B} = \mathfrak{P}_{\Lambda_0}$. But in view of the following lemma this is impossible and therefore $\mathfrak{M}_0 \neq \mathfrak{B}$.

LEMMA 16. *Let* Λ *be a linear function on* $\mathfrak{h}$ *and let* $\mathfrak{P}_\Lambda$ *be the left ideal in* $\mathfrak{B}$ *defined by* $\mathfrak{P}_\Lambda = \sum_{\beta \varepsilon P} \mathfrak{B}X_\beta + \sum_{\beta \varepsilon P} \mathfrak{B}(H_\beta - \Lambda(H_\beta))$. *Then* $\mathfrak{P}_\Lambda \neq \mathfrak{B}$.

The proof is similar to that of Lemma 5 of [5(b)]. Let us suppose $\mathfrak{P}_\Lambda = \mathfrak{B}$. Then if $\mathfrak{P} = \sum_{\beta \varepsilon P} \mathfrak{B}X_\beta$, we can choose elements $u_\beta \varepsilon \mathfrak{B}$ such that

$$1 \equiv \sum_{\beta \varepsilon P} u_\beta(H_\beta - \Lambda(H_\beta)) \mod \mathfrak{P}.$$

Again since both 1 and $H_\beta - \Lambda(H_\beta)$ are of rank zero we may assume that u_β is of rank zero. Then $u_\beta(H_\beta - \Lambda(H_\beta)) = (H_\beta - \Lambda(H_\beta))u_\beta$ and $u_\beta \varepsilon \mathfrak{B} = \mathfrak{B}\mathfrak{H}\mathfrak{U}$. Since u_β is of rank zero and we are interested only in the congruence mod $\mathfrak{P}$, we may assume that $u_\beta \varepsilon \mathfrak{H}$. Hence

$$1 - \sum_{\beta \varepsilon P} u_\beta(H_\beta - \Lambda(H_\beta)) \varepsilon \mathfrak{P} \cap \mathfrak{H} = \{0\}$$

from Lemma 33 of [5(b)]. Now $\mathfrak{H}$ is the universal enveloping algebra of the abelian Lie algebra $\mathfrak{h}$ (Lemma 21 of [5(b)]). Therefore if ρ is the representation of $\mathfrak{h}$ of degree 1 given by $\rho(H) = \Lambda(H)$ $(H \varepsilon \mathfrak{h})$, ρ can be extended uniquely to the representation of $\mathfrak{H}$ such that $\rho(1) = 1$. But since $1 = \sum_{\beta \varepsilon P} u_\beta(H_\beta - \Lambda(H_\beta))$ it follows that

$$1 = \rho(1) = \sum_{\beta \varepsilon P} \rho(u_\beta)\{\rho(H_\beta) - \Lambda(H_\beta)\} = 0$$

and so we get a contradiction. This proves the lemma.

Since $\mathfrak{M}_0 \neq \mathfrak{B}$, by Zorn's lemma[7] there exists a maximal left ideal $\mathfrak{M}$ in $\mathfrak{B}$ containing $\mathfrak{M}_0$. Let $b \rightarrow b^*$ denote the natural mapping and π the natural representation of $\mathfrak{B}$ on $\mathfrak{B}^* = \mathfrak{B}/\mathfrak{M}$. Since $\mathfrak{M}$ is maximal, π is irreducible. Moreover since $\mathfrak{M} \neq \mathfrak{B}$, $1^* \neq 0$ and $\pi(\mathfrak{M})1^* = \{0\}$. Hence if we put $V = \mathfrak{B}^*$ and $\psi = 1^*$ the first two conditions of Theorem 2 are fulfilled. So we have only to prove that $\dim \pi(\mathfrak{X})1^*$ is finite.

For any root β let s_β denote the Weyl reflexion corresponding to β and let W' be the subgroup of the Weyl group W generated by $s_i = s_{\beta_i}$, $1 \leqq i \leqq q$. Since X_i and $Y_i^{\lambda_i+1}$ $(1 \leqq i \leqq q)$ are in $\mathfrak{M}$ it follows easily (by the argument used in the proof of Lemma 6) that for any $b^* \varepsilon \mathfrak{B}^*$ we can find an integer

---

[7] Since $\mathfrak{B}$ satisfies the ascending chain conditions for left ideal (see [2]) it is not necessary to invoke Zorn's lemma.

$r\geqq 0$ such that $\pi(X_i^{r+1})b^*=\pi(Y_i^{r+1})b^*=0$, $i=1,\cdots,q$. Hence we can conclude from Lemma 1 that if Λ is a weight of π the same holds for $s\Lambda$ $(s\in W')$. Now let α be a positive compact root. Then α is not totally positive and therefore from Lemma 14, $\alpha=s\beta_i$ for some $s\in W'$ and some i $(1\leqq i\leqq q)$. Therefore $s_\alpha=ss_is^{-1}\in W'$. Now suppose d is a positive integer such that $\pi(X_{-\alpha}^{d})1^*\neq 0$. Then $\Lambda_0-d\alpha$ is a weight of π and hence $s_\alpha(\Lambda_0-d\alpha)$ $=s_\alpha\Lambda_0+d\alpha$ is also a weight of π. But then it follows from Lemma 2 that $\Lambda_0-(s_\alpha\Lambda_0+d\alpha)=(\Lambda_0(H_\alpha)-d)\alpha\geqq 0$ in $\mathfrak{F}_R$ and therefore $d\leqq\Lambda_0(H_\alpha)$. This shows that if $r>\Lambda_0(H_\alpha)$, $\pi(X_{-\alpha}^{r})1^*=0$. Hence if $\alpha_1,\cdots,\alpha_p$ are all the distinct positive compact roots of $\mathfrak{g}$, we can choose a positive integer m such that $\pi(X_{-\alpha_i}^{m})1^*=0$ $(1\leqq i\leqq p)$. From this we can conclude (see the proof of Lemma 2) that $\pi(\mathfrak{X})1^*$ is spanned by $\pi(X_{-\alpha_1}^{m_1}\cdots X_{-\alpha_p}^{m_p})1^*$, $0\leqq m_1,\cdots,m_p\leqq m$. This proves that the dimension of $\pi(\mathfrak{X})1^*$ is finite. Since the last statement of Theorem 2 is an immediate consequence of the corollary to Lemma 5, the proof of this theorem is now complete.

We denote the representation π of Theorem 2 by π_{Λ_0}. Then π_{Λ_0} is uniquely determined by Λ_0 up to equivalence.

7. **Infinitesimally unitary representations.**

For any given linear function Λ on $\mathfrak{h}$, define the left ideal $\mathfrak{P}_\Lambda$ of $\mathfrak{B}$ as in Lemma 16. Then $\mathfrak{P}_\Lambda\neq\mathfrak{B}$ and therefore by Zorn's lemma[7] there exists a maximal left ideal in $\mathfrak{B}$ containing $\mathfrak{P}_\Lambda$. Moreover we know from Lemma 5 that this maximal left ideal is unique. We shall denote it by $\mathfrak{M}_\Lambda$ and the natural representation of $\mathfrak{B}$ on $\mathfrak{B}/\mathfrak{M}_\Lambda$ by π_Λ. In case Λ satisfies the condition of Theorem 2, this definition of π_Λ agrees with the one given in Section 6. We shall say that π_Λ is *infinitesimally unitary* (see [5(c), p. 233]) if it is possible to define a scalar product (a^*,b^*) in $\mathfrak{B}^*=\mathfrak{B}/\mathfrak{M}_\Lambda$ $(a^*,b^*\in\mathfrak{B}^*)$ such that under this product $\mathfrak{B}^*$ fulfills all the requirements of a Hilbert space except completeness and $(a^*,\pi_\Lambda(X)b^*)=-(\pi_\Lambda(X)a^*,b^*)$ for all $X\in\mathfrak{g}_0$ and $a^*,b^*\in\mathfrak{B}^*$. We shall now propose to investigate the conditions on Λ under which π_Λ is infinitesimally unitary.

Let $H_1,\cdots,H_l$ be a base for $\mathfrak{h}_0$ over R and let $x_1,\cdots,x_l$ be l independent (commutative) indeterminates. Let $\mathfrak{H}$ be the subalgebra of $\mathfrak{B}$ generated by $(1,\mathfrak{h})$. Since $\mathfrak{H}$ is the universal enveloping algebra of $\mathfrak{h}$, there exists a uniquely determined homomorphism ξ_x of $\mathfrak{H}$ into $C[x]=C[x_1,\cdots,x_l]$ such that $\xi_x(H_i)=x_i$, $i=1,\cdots,l$ and $\xi_x(1)=1$. We shall now extend ξ_x to a linear mapping of $\mathfrak{B}$ into $C[x]$ as follows. Put $\mathfrak{P}=\sum_{\beta\in P}\mathfrak{B}X_\beta$ as before and let $\mathfrak{U}$ and $\mathfrak{V}$ respectively be the subalgebras of $\mathfrak{B}$ generated by $(1,X_\beta)_{\beta\in P}$ and $(1,X_{-\beta})_{\beta\in P}$. Then as we have seen above $\mathfrak{B}=\mathfrak{V}\mathfrak{H}\mathfrak{U}$ and therefore it is obvious that $\mathfrak{B} = \mathfrak{P} + \mathfrak{B}\mathfrak{H}$. Now every element of $\mathfrak{B}\mathfrak{H}$, which is of rank zero, must lie in $\mathfrak{H}$ and so for every $b \,\varepsilon\, \mathfrak{B}$ of rank zero we can choose an element $h \,\varepsilon\, \mathfrak{H}$ such that $b \equiv h \bmod \mathfrak{P}$. Moreover h is unique since $\mathfrak{P} \cap \mathfrak{H} = \{0\}$ from Lemma 33 of [5(b)]. We define $\xi_x(b) = \xi_x(h)$. Now let z be any element in $\mathfrak{B}$ and z_0 its component of rank zero (see Section 2). We extend ξ_x on $\mathfrak{B}$ by setting $\xi_x(z) = \xi_x(z_0)$. If $\mathfrak{Z}$ is the center of $\mathfrak{B}$ and χ_x is the mapping of $\mathfrak{Z}$ into $C[x]$ as defined in [5(b), Lemma 36], it is obvious that $\xi_x(z) = \chi_x(z)$ $(z \,\varepsilon\, \mathfrak{Z})$.

LEMMA 17. *Let* z_1, z_2 *be two elements of* $\mathfrak{B}$ *and suppose at least one of them has rank zero. Then* $\xi_x(z_1z_2) = \xi_x(z_2z_1) = \xi_x(z_1)\xi_x(z_2)$. *Moreover* $\xi_x(z) = \chi_x(z)$ *if* $z \,\varepsilon\, \mathfrak{Z}$.

Suppose z_1 is of rank zero and $(z_2)_0$ is the component of z_2 of rank zero. Then $z_1(z_2)_0$ and $(z_2)_0z_1$ are the components of rank zero of z_1z_2 and z_2z_1 respectively. Let $z_1 = h_1 + p_1$, $(z_2)_0 = h_2 + p_2$ where $h_1, h_2 \,\varepsilon\, \mathfrak{H}$ and $p_1, p_2 \,\varepsilon\, \mathfrak{P}$. Then $(z_2)_0z_1 \equiv z_1(z_2)_0 \equiv h_1h_2 \bmod \mathfrak{P}$ and this proves our assertion.

Λ being any linear function on $\mathfrak{h}$, we denote by $\xi_\Lambda(b)$ $(b \,\varepsilon\, \mathfrak{B})$ the value of the polynomial $\xi_x(b)$ under the substitution $x_i = \Lambda(H_i)$, $1 \leqq i \leqq l$. Let η be the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$. We can extend $-\eta$ uniquely to an anti-automorphism ϕ of $\mathfrak{B}$ over R. Then $\phi(ab) = \phi(ba)$ $(a, b \,\varepsilon\, \mathfrak{B})$ and $\phi(X) = -\eta(X)$ $(X \,\varepsilon\, \mathfrak{g})$.

LEMMA 18. *Let* Λ *be a linear function on* $\mathfrak{h}$. *Then* π_Λ *is infinitesimally unitary if and only if* $\xi_\Lambda(\phi(z)z)$ *is real and nonnegative for every* $z \,\varepsilon\, \mathfrak{B}$.

Suppose $\pi = \pi_\Lambda$ is infinitesimally unitary. Let $a \rightarrow a^*$ denote the natural mapping of $\mathfrak{B}$ on $\mathfrak{B}^* = \mathfrak{B}/\mathfrak{M}_\Lambda$ and (a^*, b^*) the corresponding scalar product on $\mathfrak{B}^*$. We assume, as we may, that $(1^*, 1^*) = 1$. Then if $H \,\varepsilon\, \mathfrak{h}_0$, $(1^*, \pi(H)1^*) = -(\pi(H)1^*, 1^*)$ and therefore $\Lambda(H)$ is purely imaginary. Now suppose z is an element of $\mathfrak{B}$ of rank λ so that $[H, z] = \lambda(H)z$ $(H \,\varepsilon\, \mathfrak{h})$. Then if $H \,\varepsilon\, \mathfrak{h}_0$,

$$\begin{aligned}\lambda(H)(1^*, \pi(z)1^*) &= (1^*, \pi([H, z])1^*) \\ &= -(\pi(H)1^*, \pi(z)1^*) - (1^*, \pi(zH)1^*) \\ &= \Lambda(H)(1^*, \pi(z)1^*) - \Lambda(H)(1^*, \pi(z)1^*) = 0\end{aligned}$$

and so if $\lambda \neq 0$, $(1^*, \pi(z)1^*) = 0$. Therefore if b is any element in $\mathfrak{B}$ and b_0 is its component of rank zero, $(1^*, \pi(b)1^*) = (1^*, \pi(b_0)1^*)$. Moreover if $b_0 \equiv h \bmod \mathfrak{P}$ $(h \,\varepsilon\, \mathfrak{H})$, it is obvious that

$$(1^*, \pi(b)1^*) = (1^*, \pi(h)1^*) = \xi_\Lambda(h) = \xi_\Lambda(b)$$

since $\pi(\mathfrak{P})1^* = \{0\}$. This proves that $(1^*, \pi(b)1^*) = \xi_\Lambda(b)$ for all $b \in \mathfrak{B}$. On the other hand $(\pi(X)a^*, b^*) = (a^*, \pi(\phi(X))b^*)$ $(a^*, b^* \in \mathfrak{B}, X \in \mathfrak{g})$ and so it is clear that $(\pi(z)a^*, b^*) = (a^*, \pi(\phi(z)b^*)$ $(z \in \mathfrak{B})$. Hence

$$\xi_\Lambda(\phi(z)z) = (1^*, \pi(\phi(z)z)1^*) = (\pi(z)1^*, \pi(z)1^*) \geqq 0.$$

This shows that our condition is necessary.

Conversely suppose $\xi_\Lambda(\phi(z)z)$ is real and nonnegative for all $z \in \mathfrak{B}$. Put $(w, z) = \xi_\Lambda(\phi(w)z)$ $(w, z \in \mathfrak{B})$ and $|z| = (z, z)^{\frac{1}{2}} \geqq 0$. Then a well-known elementary argument shows that $|(w, z)| \leqq |w|\,|z|$ and $|z + w| \leqq |z| + |w|$. Let $\mathfrak{M}$ be the set of all elements $z \in \mathfrak{B}$ such that $|z| = 0$. Since $|z_1 + z_2| \leqq |z_1| + |z_2|$, $\mathfrak{M}$ is a vector space. Moreover if b, $z, w \in \mathfrak{B}$ and $b' = \phi(b)$,

$$(w, bz) = \xi_\Lambda(\phi(w)bz) = \xi_\Lambda(\phi(b'w)z) = (b'w, z).$$

Hence if $|z| = 0$, $|(w, bz)| \leqq |b'w|\,|z| = 0$ and therefore in particular $|bz|^2 = 0$. This proves that $\mathfrak{M}$ is a left ideal. Let $b \rightarrow b^*$ denote the natural mapping and π the natural representation of $\mathfrak{B}$ on $\mathfrak{B}^* = \mathfrak{B}/\mathfrak{M}$. It is clear that (w, z) depends only on w^* and z^*. Put $(w^*, z^*) = (w, z)$ and $|z^*| = |z|$. Then $|z^*| \neq 0$ unless $z^* = 0$ and if $X \in \mathfrak{g}_0$,

$$\begin{aligned}(w^*, \pi(X)z^*) = (w, Xz) = \xi_\Lambda(\phi(w)Xz) &= -\xi_\Lambda(\phi(Xw)z) \\ = -(\pi(X)w^*, z^*) & \qquad\qquad (w, z \in \mathfrak{B}).\end{aligned}$$

Hence in order to prove the lemma it only remains to prove that $\mathfrak{M} = \mathfrak{M}_\Lambda$. It is clear that $\xi_\Lambda(z) = 0$ if $z \in \mathfrak{P}$ and therefore it follows from Lemma 17 that $\xi_\Lambda(\mathfrak{P}_\Lambda) = \{0\}$. But since $\mathfrak{P}_\Lambda$ is a left ideal this implies that $\mathfrak{M} \supset \mathfrak{P}_\Lambda$. Hence it would be enough to prove that $\mathfrak{M}$ is maximal. Since $\xi_\Lambda(1) = 1 \neq 0$, $\mathfrak{M} \neq \mathfrak{B}$. Now suppose $\mathfrak{M}'$ is a left ideal in $\mathfrak{B}$ such that $\mathfrak{B} \neq \mathfrak{M}'$ and $\mathfrak{M}' \supset \mathfrak{M}$. Then if $b' \in \mathfrak{M}'$ and $H \in \mathfrak{h}$

$$[H, b'] = Hb' - b'(H - \Lambda(H)) - \Lambda(H)b' \in \mathfrak{M}'$$

and from this we can conclude that all the components of b' of various ranks also lie in $\mathfrak{M}'$. Let $(b')_0$ be its component of rank zero and choose $h \in \mathfrak{H}$ such that $(b')_0 \equiv h \bmod \mathfrak{P}$. Then it is obvious that $(b')_0 \equiv h \equiv \xi_\Lambda(h) \bmod \mathfrak{P}_\Lambda$. Since $(b')_0$ and $\mathfrak{P}_\Lambda$ are both contained in $\mathfrak{M}'$, $\xi_\Lambda(b') = \xi_\Lambda(h) \in \mathfrak{M}'$. However $\mathfrak{M}' \neq \mathfrak{B}$ and so this implies that $\xi_\Lambda(b') = 0$. Now put $b' = \phi(z')z'$ where z' is any element in $\mathfrak{M}'$. Then $b' \in \mathfrak{M}'$ and therefore $(z', z') = \xi_\Lambda(\phi(z')z') = 0$. This shows that $z' \in \mathfrak{M}$ and hence $\mathfrak{M}' = \mathfrak{M}$. Our proof is now complete.

Notice that it follows automatically from the above condition on ξ_Λ that $\dim \pi(\mathfrak{X})1^*$ is finite. For this it would be enough to verify (see Lemma 2) that if α is a positive compact root of $\mathfrak{g}$, $\pi(X_{-\alpha}{}^r)1^* = 0$ for some integer $r > 0$. But $\eta(X_\alpha) = -X_{-\alpha}$ (see Section 4) and therefore

$$\xi_\Lambda(\phi(X_{-\alpha}{}^r)X_{-\alpha}{}^r) = \xi_\Lambda(X_\alpha{}^r X_{-\alpha}{}^r) \geqq 0 \qquad (r \geqq 1).$$

Moreover we know from Lemma 6 of [5(b)] that

$$\begin{aligned} X_\alpha{}^r X_\alpha{}^{-r} &= X_\alpha{}^{r-1}[X_\alpha, X_{-\alpha}{}^r] + X_\alpha{}^{r-1} X_{-\alpha}{}^r X_\alpha \\ &\equiv r X_\alpha{}^{r-1} X_{-\alpha}{}^{r-1}(\Lambda(H_\alpha) - r + 1) \bmod \mathfrak{P}_\Lambda \qquad (r \geqq 1). \end{aligned}$$

Therefore if $\lambda_\alpha = \Lambda(H_\alpha)$,

$$0 \leqq \xi_\Lambda(X_\alpha{}^r X_{-\alpha}{}^r) = r! \prod_{1 \leqq j \leqq r} (\lambda_\alpha - j + 1)$$

by induction on r. Putting $r = 1$ we conclude that $\lambda_\alpha \geqq 0$. Now let r be the least integer such that $\lambda_\alpha - r + 1 \leqq 0$. Then the above product is negative unless $\lambda_\alpha = r - 1$. Hence we conclude that λ_α is a nonnegative integer and therefore $\xi_\Lambda(X_\alpha{}^r X_{-\alpha}{}^r) = 0$ for $r \geqq \lambda_\alpha + 1$. But since

$$\xi_\Lambda(X_\alpha{}^r X_{-\alpha}{}^r) = |\pi(X_{-\alpha}{}^r)1^*|^2,$$

$\pi(X_{-\alpha}{}^r)1^* = 0$ for $r \geqq \lambda_\alpha + 1$. This shows that π then satisfies all the requirements of Theorem 1 and therefore $\Lambda(H_\beta)$ must be a nonnegative integer for all positive roots β which are not totally positive.

Now suppose β is a noncompact root. Then $\eta(X_{-\beta}) = X_\beta$ (see Section 4) and therefore

$$\xi_\Lambda(\phi(X_{-\beta}{}^r)X_{-\beta}{}^r) = (-1)^r \xi_\Lambda(X_\beta{}^r X_{-\beta}{}^r) = r! \prod_{1 \leqq j \leqq r} (j - 1 - \lambda_\beta) \quad (r \geqq 1)$$

where $\lambda_\beta = \Lambda(H_\beta)$. Since again this product must always be real and nonnegative, we conclude that λ_β is real and $\leqq 0$. In particular if β is not totally positive it follows, in view of our earlier remarks, that $\lambda_\beta = 0$. Thus we have obtained the following corollary.

COROLLARY 1. *In order that* π_Λ *be infinitesimally unitary it is necessary that* Λ *should satisfy the following conditions:*

(1) $\Lambda(H_\alpha)$ *is a nonnegative integer for every positive compact root* α.

(2) $\Lambda(H_\beta) = 0$ *for every noncompact positive root* β *which is not totally positive.*

(3) $\Lambda(H_\gamma)$ *is real and* $\leqq 0$ *for every totally positive root* γ.

The space $\mathfrak{F}$ of linear functions on $\mathfrak{h}$, being a finite-dimensional vector space over C, has a natural topology.

COROLLARY 2. *The set of all linear functions* Λ *on* $\mathfrak{h}$ *for which* π_Λ *is infinitesimally unitary, is closed in* $\mathfrak{F}$.

For a fixed element $b \varepsilon \mathfrak{B}$, $\Lambda \rightarrow \xi_\Lambda(b)$ $(\Lambda \varepsilon \mathfrak{F})$ is obviously a polynomial function on $\mathfrak{F}$. Hence it is continuous and therefore the corollary follows from Lemma 18.

Define $\mathfrak{g}'$ as in the proof of Lemma 13.

LEMMA 19. *Suppose* Λ *satisfies the conditions of Theorem* 2 *and* $\Lambda(H_\beta) = 0$ *for every positive root* β *which is not totally positive. Then* $\pi_\Lambda(X) = 0$ *for all* $X \varepsilon \mathfrak{g}'$.

We use the notation of the proof of Lemma 13. Let α be a positive compact root of $\mathfrak{g}'$. Then we can choose a noncompact root β of $\mathfrak{g}'$ such that $\beta + \alpha$ is a root. For otherwise $[X_\alpha, X_\beta] = 0$ for every noncompact root β of $\mathfrak{g}'$. Since it is obvious that $\mathfrak{q} = \mathfrak{g}' \cap \mathfrak{p}$ is spanned by the elements X_β corresponding to all noncompact roots β of $\mathfrak{g}'$, X_α lies in the centralizer $\mathfrak{l}$ of $\mathfrak{q}$ in $\mathfrak{k}$. But $\mathfrak{l} \subset \mathfrak{g}_+$ and therefore $X_\alpha \varepsilon \mathfrak{g}_+ \cap \mathfrak{g}' = \{0\}$, which of course is false. So let β be a noncompact root of $\mathfrak{g}'$ such that $\gamma = \beta + \alpha$ is a root. Since $\mathfrak{q} \cap (\mathfrak{p}_+ + \mathfrak{p}_-) = \{0\}$, no root of $\mathfrak{g}'$ can be totally positive or totally negative. Therefore $\Lambda(H_\beta) = \Lambda(H_\gamma) = 0$. But since $\alpha = \gamma - \beta$, H_α is a linear combination of H_β and H_γ and so $\Lambda(H_\alpha) = 0$. This shows that $\Lambda(H_\beta) = 0$ for every root β of $\mathfrak{g}'$. Moreover if β is a positive root of $\mathfrak{g}'$, it follows from Theorem 1 that $\pi_\Lambda(X_{-\beta})1^* = 0$ since β is not totally positive. Therefore $\pi_\Lambda(\mathfrak{g}')1^* = \{0\}$ and since $\mathfrak{g}'$ is an ideal in $\mathfrak{g}$ and π_Λ is irreducible, $\pi_\Lambda(\mathfrak{g}') = \{0\}$.

COROLLARY 1. *If* π_Λ *is infinitesimally unitary* $\pi_\Lambda(\mathfrak{g}') = \{0\}$.

This is an immediate consequence of the above lemma and Corollary 1 to Lemma 18.

COROLLARY 2. *Suppose* $\mathfrak{g}_0$ *is simple and* $\mathfrak{g}_0 \neq \mathfrak{k}_0$. *Then if* $\mathfrak{c}_0 = \{0\}$, π_Λ *cannot be infinitesimally unitary unless* $\Lambda = 0$ *in which case it is the zero representation of* $\mathfrak{g}$ *of degree* 1.

For then we know that $\mathfrak{g} = \mathfrak{g}'$ (see the proof of Corollary 2 to Lemma 13). Now apply Corollary 1 above.

In Corollary 1 to Lemma 18 we have obtained certain necessary conditions on Λ in order that π_Λ may be infinitesimally unitary. It is natural to ask how far these conditions are also sufficient. The following result gives a partial answer to this question.

THEOREM 3. *Let 2ρ denote the sum of all positive roots of $\mathfrak{g}$ and let Λ be a linear function on $\mathfrak{h}$ satisfying the following conditions:*

(1) $\Lambda(H_\alpha)$ *is a nonnegative integer for every positive compact root* α.

(2) $\Lambda(H_\beta)=0$ *for every noncompact positive root* β *which is not totally positive.*

(3) $\Lambda(H_\gamma)+\rho(H_\gamma)$ *is real and* $\leqq 0$ *for every totally positive root* γ. *Then* π_Λ *is infinitesimally unitary.*

We leave the proof of this theorem to a subsequent paper (see however [5(h)]).

8. Representations of the group. Let G be the simply connected Lie group with the Lie algebra $\mathfrak{g}_0$ and let K be its analytic subgroup corresponding to $\mathfrak{k}_0$. Let Λ be a linear function on $\mathfrak{h}$ which fulfills the conditions of Theorem 2. We shall now show that π_Λ can be 'extended' to a representation of G on a Hilbert space. Let Ω be the set of all equivalence classes of finite-dimensional irreducible representations of K. Henceforth we shall make free use of the terminology of [5(c)].

THEOREM 4. *Let Λ be a linear function on $\mathfrak{h}$ such that $\Lambda(H_\beta)$ is a nonnegative integer for every positive root β which is not totally positive. Then there exists an irreducible quasi-simple representation σ of G on a Hilbert space $\mathfrak{H}$ with the following property. Let $\mathfrak{H}_{\mathfrak{D}}$ $(\mathfrak{D}\,\varepsilon\,\Omega)$ denote the subspace of $\mathfrak{H}$ consisting of all those elements which transform under $\sigma(K)$ according to $\mathfrak{D}$. Then the corresopnding representation σ_0 of $\mathfrak{B}$ defined on $\mathfrak{H}_0=\sum_{\mathfrak{D}\varepsilon\Omega}\mathfrak{H}_{\mathfrak{D}}$ is equivalent to π_Λ.*

Put $\pi=\pi_\Lambda$ and define $\mathfrak{P}_\Lambda$ and $\mathfrak{M}_\Lambda$ as in Section 7. If z is any element in $\mathfrak{B}$ of rank zero, $z\equiv\xi_\Lambda(z) \bmod \mathfrak{P}_\Lambda$ and therefore in particular if $z\,\varepsilon\,\mathfrak{Z}$, $z-\xi_\Lambda(z)\,\varepsilon\,\mathfrak{M}_\Lambda$. Moreover we know from Theorem 2 that $\dim \mathfrak{X}/\mathfrak{X}\cap\mathfrak{M}_\Lambda$ is finite and so we can conclude from Lemma 9 of [5(c)] that π is a quasi-simple[8] irreducible representation of $\mathfrak{B}$.

Let $(\beta_1,\cdots,\beta_l)$ and $(\beta_1,\cdots,\beta_q)$ be defined as in the proof of Theorem 2 and let Λ' be the linear function on $\mathfrak{h}$ given by $\Lambda'(H_{\beta_i})=\Lambda(H_{\beta_i})$, $1\leqq i\leqq q$ and $\Lambda'(H_{\beta_i})=0$, $q<i\leqq l$. Then from Theorem 1 of [5(b)], there exists a finite-dimensional irreducible representation ρ of $\mathfrak{g}$ with the highest weight Λ'. Let U be the representation space of ρ and ϕ a nonzero vector in U belonging to the weight Λ'. Then if α is any positive compact

---

[8] See [5(d), p. 56] for the definition of quasi-simplicity in this case.

root of $\mathfrak{g}$, $\rho(X_\alpha)\phi = 0$ and $\rho(H_\alpha)\phi = \Lambda'(H_\alpha)\phi = \Lambda(H_\alpha)\phi$ since α is a linear combination of $\beta_1, \cdots, \beta_q$ (Lemma 13). Similarly if $\psi \neq 0$ is a vector belonging to the weight Λ of π, $\pi(X_\alpha)\psi = 0$ and $\pi(H_\alpha)\psi = \Lambda(H_\alpha)\psi$. Let $\mathfrak{X}'$ be the subalgebra of $\mathfrak{B}$ generated by $(1\,\mathfrak{k}')$. Since $\mathfrak{k}'$ is semisimple we can apply Lemma 2 to $\mathfrak{X}'$ and conclude that the representations of $\mathfrak{k}'$ defined on $\pi(\mathfrak{X}')\psi$ and $\rho(\mathfrak{X}')\phi$ are both irreducible and therefore from Lemma 5 they are equivalent. Since ρ is a finite-dimensional representation of $\mathfrak{g}$, the existence of σ now follows from Theorem 4 of [5(d)].

COROLLARY. *Let* λ, μ *be two linear functions on* $\mathfrak{h}$. *Then if* π_λ *and* π_μ *are both infinitesimally unitary the same holds for* $\pi_{\lambda+\mu}$.

Suppose π_λ and π_μ are infinitesimally unitary. Then λ and μ both satisfy the conditions of Theorem 2 (see Corollary 1 to Lemma 18) and therefore from Theorem 4 above and Theorem 9 of [5(c)] we can find two irreducible unitary representations σ_λ, σ_μ of G on the Hilbert spaces $\mathfrak{H}_\lambda$, $\mathfrak{H}_\mu$ respectively such that the corresponding representations of $\mathfrak{B}$ defined on $\mathfrak{H}_\lambda{}^0 = \sum_{\mathfrak{D}\varepsilon\Omega} (\mathfrak{H}_\lambda)_{\mathfrak{D}}$ and $\mathfrak{H}_\mu{}^0 = \sum_{\mathfrak{D}\varepsilon\Omega} (\mathfrak{H}_\mu)_{\mathfrak{D}}$ are equivalent to π_λ and π_μ. We denote these representations of $\mathfrak{B}$ by $\sigma_\lambda{}^0$ and $\sigma_\mu{}^0$. Then it is clear from the proofs of Lemma 18 and Theorem 9 of [5(c)] that we can choose $\psi_\lambda \varepsilon \mathfrak{H}_\lambda{}^0$ and $\psi_\mu \varepsilon \mathfrak{H}_\mu{}^0$ such that

$$(\psi_\lambda, \sigma_\lambda{}^0(z)\psi_\lambda) = \xi_\lambda(z), \qquad (\psi_\mu, \sigma_\mu{}^0(z)\psi_\mu) = \xi_\mu(z)$$

for all $z \varepsilon \mathfrak{B}$. Let $\mathfrak{H}$ be the Hilbert space which is the Kronecker product of $\mathfrak{H}_\lambda$ and $\mathfrak{H}_\mu$ (in the sense of Hilbert space theory) and let $\psi = \psi_\lambda \times \psi_\mu$. Then if ψ' and ψ'' are well-behaved under σ_λ and σ_μ respectively it is obvious that $\psi' \times \psi''$ is well-behaved under the representation $\sigma = \sigma_\lambda \times \sigma_\mu$ of G on $\mathfrak{H}$. Let σ^0 denote the representation of $\mathfrak{B}$ defined under σ on the subspace $\mathfrak{H}^0$ consisting of all well-behaved elements in $\mathfrak{H}$. Then it is obvious that $\sigma^0(X)(\phi_1 \times \phi_2) = (\sigma_\lambda{}^0(X)\phi_1) \times \phi_2 + \phi_1 \times (\sigma_\mu{}^0(X)\phi_2)$ $(X \varepsilon \mathfrak{g}, \phi_1 \varepsilon \mathfrak{H}_\lambda{}^0, \phi_2 \varepsilon \mathfrak{H}_\mu{}^0)$ and therefore $\sigma^0(X_\beta)\psi = 0$ for all positive roots β and $\sigma^0(H)\psi = (\lambda(H) + \mu(H))\psi$ if $H \varepsilon \mathfrak{h}$. Hence $\sigma^0(\mathfrak{P}_{\lambda+\mu})\psi = \{0\}$ in the notation of Section 7. On the other hand since σ is a unitary representation of G,

$$(\phi', \sigma^0(X)\phi) = -(\sigma^0(X)\phi', \phi) \qquad (X \varepsilon \mathfrak{g}_0; \phi, \phi' \varepsilon \mathfrak{H}^0).$$

Therefore if z is any element in $\mathfrak{B}$ of rank $\nu \neq 0$ it follows (see the proof of Lemma 18) that $(\psi, \sigma^0(z)\psi) = 0$. This shows that

$$(\psi, \sigma^0(b)\psi) = \xi_{\lambda+\mu}(b)\,|\,\psi\,|^2 \qquad (b \varepsilon \mathfrak{B})$$

and therefore

$$|\sigma^0(z)\psi|^2=(\sigma^0(z)\psi,\sigma^0(z)\psi)=(\psi,\sigma^0(\phi(z)z)\psi)=\xi_{\lambda+\mu}(\phi(z)z)\,|\psi|^2$$

for any $z\in\mathfrak{B}$. Hence $\pi_{\lambda+\mu}$ is infinitesimally unitary from Lemma 18.

9. An auxiliary result.

In this section we shall prove a result which will be required in another paper. Let Λ_0 be a linear function on $\mathfrak{h}$ satisfying the conditions of Theorem 3. As before let P be the set of all positive roots and P_k and P_+ respectively the subsets of P consisting of all compact and all totally positive roots. Let P' be the complement of $P_k\cup P_+$ in P. We know (Lemma 5) that there exists a unique left ideal $\mathfrak{Y}_{\Lambda_0}$ in $\mathfrak{X}$ containing

$$\sum_{\alpha\varepsilon P_k}\mathfrak{X}X_\alpha+\sum_{\gamma\varepsilon P}\mathfrak{X}(H_\gamma-\Lambda_0(H_\gamma)).$$

Our object now is to prove the following theorem.

THEOREM 5. *Let* Λ_0 *be a linear function on* $\mathfrak{h}$ *satisfying the conditions of Theorem* 3. *Then*

$$\mathfrak{M}=\sum_{\gamma\varepsilon P}\mathfrak{B}X_\gamma+\sum_{\gamma\varepsilon P'}\mathfrak{B}X_{-\gamma}+\mathfrak{B}\mathfrak{Y}_{\Lambda_0}$$

is a maximal left ideal in $\mathfrak{B}$.

It is obvious that Λ_0 fulfills the conditions of Theorem 2 and therefore we can consider the representation π_{Λ_0} and choose a vector $\psi\neq 0$ in the representation space corresponding to Theorem 2. Define $\mathfrak{g}_+$ and $\mathfrak{g}'$ as in Section 5. Then we know from Lemma 19 that $\pi_{\Lambda_0}(\mathfrak{g}')=\{0\}$. Moreover it follows from Lemmas 2 and 5 (applied to $\mathfrak{k}$) that $\pi_{\Lambda_0}(\mathfrak{Y}_{\Lambda_0})\psi=\{0\}$. This shows that $\pi_{\Lambda_0}(\mathfrak{M})\psi=\{0\}$ and therefore $\mathfrak{M}\neq\mathfrak{B}$ since $\psi\neq 0$.

Let $\mathfrak{M}_1$ be the set of all elements $b\,\varepsilon\,\mathfrak{B}$ such that $\pi_{\Lambda_0}(b)\psi=0$. Since $\mathfrak{Y}_{\Lambda_0}$ is a maximal left ideal in $\mathfrak{X}$, it is clear that $\mathfrak{M}_1\cap\mathfrak{X}=\mathfrak{Y}_{\Lambda_0}$. Therefore in view of the fact that $\mathfrak{g}'\subset\mathfrak{M}_1$, we can conclude that $\mathfrak{g}'\cap\mathfrak{k}\subset\mathfrak{Y}_{\Lambda_0}$. On the other hand $\mathfrak{g}'=\sum_{\gamma\varepsilon P'}(CX_\gamma+CX_{-\gamma})+\mathfrak{g}'\cap\mathfrak{k}$ and so it follows that $\mathfrak{g}'\subset\mathfrak{M}$.

Let $b\to b^*$ $(b\,\varepsilon\,\mathfrak{B})$ denote the natural mapping and π the natural representation of $\mathfrak{B}$ on $\mathfrak{B}^*=\mathfrak{B}/\mathfrak{M}$. In order to prove that $\mathfrak{M}$ is a maximal left ideal in $\mathfrak{B}$ it is sufficient to show that π is irreducible. Let $\mathfrak{B}_+$ be the subalgebra of $\mathfrak{B}$ generated by $(1,\mathfrak{g}_+)$. Since $\mathfrak{g}'$ is an ideal in $\mathfrak{g}$ and $\mathfrak{g}'\subset\mathfrak{M}$, it follows that $\pi(\mathfrak{g}')=\{0\}$ and therefore $\mathfrak{B}^*=\pi(\mathfrak{B}_+)1^*$. Hence it would be enough to prove that $\mathfrak{B}_+\cap\mathfrak{M}$ is a maximal left ideal in $\mathfrak{B}_+$. Let $\mathfrak{X}_+$ be the subalgebra of $\mathfrak{X}$ generated by $(1,\mathfrak{k}\cap\mathfrak{g}_+)$. Since $\pi(\mathfrak{Y}_{\Lambda_0})1^*=\{0\}$ and since $\mathfrak{Y}_{\Lambda_0}$ is a maximal left ideal in $\mathfrak{X}$, it is clear that $\mathfrak{X}^*=\pi(\mathfrak{X})1^*$ is irreducible under $\pi(\mathfrak{X})$ and therefore also under $\pi(\mathfrak{X}_+)$. Hence $\mathfrak{Y}_+ = \mathfrak{X}_+ \cap \mathfrak{Y}_{\Lambda_0}$ is a maximal left ideal in $\mathfrak{X}_+$. Let Q denote the set of all positive roots of $\mathfrak{g}_+$ (see Section 5). Then from Lemma 5, $\mathfrak{Y}_+$ must be the unique maximal left ideal in $\mathfrak{X}_+$ containing

$$\sum_{\gamma \varepsilon Q_k} \mathfrak{X}_+ X_\alpha + \sum_{\gamma \varepsilon Q} \mathfrak{X}_+ (H_\gamma - \Lambda_0(H_\gamma))$$

where $Q_k = Q \cap P_k$. Now if $\mathfrak{M}_+ = \sum_{\gamma \varepsilon Q} \mathfrak{B}_+ X_\gamma + \mathfrak{B}_+ \mathfrak{Y}_+$ it is obvious that $\mathfrak{B}_+ \cap \mathfrak{M} \supset \mathfrak{M}_+$ and therefore it is enough to prove that $\mathfrak{M}_+$ is a maximal left ideal in $\mathfrak{B}_+$. Let $2\rho_+ = \sum_{\gamma \varepsilon Q} \gamma$ and $2\rho = \sum_{\gamma \varepsilon P} \gamma$. If β is a root in $\mathfrak{g}'$ and γ a root in $\mathfrak{g}_+$ then $\beta(H_\gamma) = 0$. Hence it follows that $\rho(H_\gamma) = \rho_+(H_\gamma)$ for all $\gamma \varepsilon Q$. This shows (by comparison with the conditions of Theorem 3) that the problem of proving the maximality of $\mathfrak{M}_+$ in $\mathfrak{B}_+$ is exactly the same as our original problem if we replace $\mathfrak{g}$ by $\mathfrak{g}_+$. Therefore we may assume that $\mathfrak{g}' = \{0\}$ and $P = P_k \cup P_+$.

Let $\mathfrak{M}_1 \neq \mathfrak{M}$ be a left ideal in $\mathfrak{B}$ containing $\mathfrak{M}$. We have to show that $\mathfrak{M}_1 = \mathfrak{B}$. For any linear function Λ on $\mathfrak{h}$, let $\mathfrak{B}_\Lambda{}^*$ denote the set of elements $b^* \varepsilon \mathfrak{B}^*$ such that $\pi(H) b^* = \Lambda(H) b^*$ $(H \varepsilon \mathfrak{h})$. Then it is clear that $\mathfrak{B}^* = \sum_\Lambda \mathfrak{B}_\Lambda{}^*$ and $\mathfrak{B}_\Lambda{}^* = \{0\}$ unless $\Lambda_0 - \Lambda$ is a linear combination of positive roots with coefficients which are all nonnegative integers (see the proof of Lemma 2). Moreover $\mathfrak{M}_1{}^* = \sum_\Lambda (\mathfrak{M}_1{}^* \cap \mathfrak{B}_\Lambda{}^*) \neq \{0\}$ since $\mathfrak{M}_1 \neq \mathfrak{M}$ (Lemma1 of [5(b)]). Hence in view of the above remarks, there exists a highest weight Λ_1 such that $\mathfrak{M}_1{}^* \cap \mathfrak{B}_{\Lambda_1}{}^* \neq \{0\}$. Choose an element $\phi \neq 0$ in $\mathfrak{M}_1{}^* \cap \mathfrak{B}_{\Lambda_1}{}^*$. Then if γ is a positive root $\mathfrak{M}_1{}^* \cap \mathfrak{B}_{\Lambda_1+\gamma}{}^* = \{0\}$ and therefore $\pi(X_\gamma)\phi = 0$. Now define $\chi_x(z)$ $(z \varepsilon \mathfrak{Z})$ as in Lemma 17 and for any linear function λ on $\mathfrak{h}$ let $\chi_\lambda(z)$ denote the value of this polynomial under the substitution $x_i = \lambda(H_i)$ $(1 \leqq i \leqq l)$. (Here we are using the notation of Section 7). Then it is clear that $\pi(z)\phi = \chi_{\Lambda_1}(z)\phi$ for all $z \varepsilon \mathfrak{Z}$. Similarly $\pi(z) 1^* = \chi_{\Lambda_0}(z) 1^*$ $(z \varepsilon \mathfrak{Z})$ and therefore $\pi(z) b^* = \chi_{\Lambda_0}(z) b^*$ for every $b^* \varepsilon \mathfrak{B}^*$. Since $\phi \neq 0$, we conclude that $\chi_{\Lambda_0}(z) = \chi_{\Lambda_1}(z)$ for any $z \varepsilon \mathfrak{Z}$. But then if W is the Weyl group of $\mathfrak{g}$ with respect to $\mathfrak{h}$, it follows from Theorem 5 of [5(b)] that $\Lambda_1 + \rho = s_0(\Lambda_0 + \rho)$ for some $s_0 \varepsilon W$. (Here $2\rho = \sum_{\gamma \varepsilon P} \gamma$).

Let N_- denote the set of all totally negative roots (Section 4) and put $\bar{P} = P_k \cup N_-$. Since $\mathfrak{g}' = \{0\}$, it follows from Lemmas 11 and 12 that $\bar{P}$ satisfies the conditions of Corollary 1 to Lemma 4. Put $\Lambda_0' = \Lambda_0 + \rho$ and let $\mathfrak{F}_R(\bar{P})$ denote the set of those linear functions on $\mathfrak{h}$ which are linear combinations of roots in $\bar{P}$ with real nonnegative coefficients. Since

$\rho(H_\beta) > 0$ for every $\beta \varepsilon P$ (see Weyl [7]) it follows from our assumptions on Λ_0 that $\Lambda_0'(H_\gamma) \geqq 0$ for all $\gamma \varepsilon \bar{P}$. As we shall see presently (Lemma 20 below), this implies that $\Lambda_0' - s\Lambda_0' \varepsilon \mathfrak{F}^+(\bar{P})$ for every $s \varepsilon W$. Let $(\beta_1, \cdots, \beta_l)$ be a fundamental system of roots in P. Suppose $(\beta_1, \cdots, \beta_t)$ are all the totally positive roots and $(\alpha_1, \cdots, \alpha_r)$ all the (distinct) compact roots among $(\beta_1, \cdots, \beta_l)$. Then we know from Lemma 13 that every totally negative root is of the form $-(\beta_i + m_1\alpha_1 + \cdots + m_r\alpha_r)$ where $1 \leqq i \leqq t$ and $m_1, \cdots, m_r \geqq 0$. Hence

$$\Lambda_0 - \Lambda_1 = (\Lambda_0 + \rho) - s_0(\Lambda_0 + \rho)$$
$$= -(\mu_1\beta_1 + \cdots + \mu_t\beta_t) + m_1'\alpha_1 + \cdots + m_r'\alpha_r$$

where $\mu_1, \cdots, \mu_t \geqq 0$. On the other hand $\mathfrak{B}_{\Lambda_1}^* \neq \{0\}$ and therefore, as we have seen above,

$$\Lambda_0 - \Lambda_1 = \nu_1\beta_1 + \cdots + \nu_t\beta_t + n_1\alpha_1 + \cdots + n_r\alpha_r$$

where $\nu_1, \cdots, \nu_t, n_1, \cdots, n_r \geqq 0$. Since $(\beta_1, \cdots, \beta_l)$ are linearly independent we conclude that $-\mu_i = \nu_i = 0$ $(1 \leqq i \leqq t)$ and therefore $\Lambda_0 - \Lambda_1$ is a linear combination of compact roots. Let $\gamma_1, \cdots, \gamma_p$ be all the (distinct) totally positive roots and $\delta_1, \cdots, \delta_q$ all the compact roots in P. Then $\mathfrak{B}^*$ is spanned by elements of the form $\pi(X_{-\gamma_1}^{\mu_1} \cdots X_{-\gamma_p}^{\mu_p} X_{-\delta_1}^{\lambda_1} \cdots X_{-\delta_1}^{\lambda_q})1^*$ where λ_i and μ_i are nonnegative integers (see the proof of Lemma 2). Since $\phi \varepsilon \mathfrak{B}_{\Lambda_1}^*$, it must be a linear combination of those elements of the above form for which

$$\mu_1\gamma_1 + \cdots + \mu_p\gamma_p + \lambda_1\delta_1 + \cdots + \lambda_q\delta_q = \Lambda_0 - \Lambda_1.$$

However since $\Lambda_0 - \Lambda_1$ is a linear combination of compact roots it follows from Lemma 13 that $\mu_1 = \cdots = \mu_p = 0$ and therefore $\phi \varepsilon \pi(\mathfrak{X})1^* = \mathfrak{X}^*$. But $\pi(X_\alpha)\phi = 0$ $(\alpha \varepsilon P_k)$ and as we have seen above $\mathfrak{X}^*$ is irreducible under $\pi(\mathfrak{X})$. Therefore it follows from Lemma 2 (applied to $\mathfrak{X}$ and ϕ) that $\Lambda_1 - \Lambda_0 \geqq 0$. But $\Lambda_0 - \Lambda_1 = \lambda_1\delta_1 + \cdots + \lambda_q\delta_q \geqq 0$. Hence $\Lambda_0 = \Lambda_1$ and $\lambda_1 = \cdots = \lambda_q = 0$. This shows that $\phi = c1^*$ $(c \varepsilon C)$ and since $\phi \neq 0$ it follows that $\mathfrak{M}_1^* \supset \pi(\mathfrak{B})\phi = \mathfrak{B}^*$ and therefore $\mathfrak{M}_1 = \mathfrak{B}$. Hence in order to complete the proof it remains to prove the following lemma.

LEMMA 20. *Let $\bar{P}_0$ be a fundamental system of roots of $\mathfrak{g}$ and let $\mathfrak{F}^+(\bar{P}_0)$ denote the set of all linear functions on $\mathfrak{h}$ which can be written as linear combinations of roots in $\bar{P}_0$ with real nonnegative coefficients. Then if $\Lambda \varepsilon \mathfrak{F}_R$ and $\Lambda(H_\beta) \geqq 0$ for all $\beta \varepsilon \bar{P}_0$, $\Lambda - s\Lambda \varepsilon \mathfrak{F}^+(\bar{P}_0)$ for every $s \varepsilon W$.*

It is enough to prove this lemma under the additional assumption that

$\Lambda(H_\beta)$ $(\beta \varepsilon P_0)$ are all rational since the general case then follows immediately by continuity. Moreover by multiplying Λ by a suitable positive integer, the problem may be reduced to the case when $\Lambda(H_\beta)$ $(\beta \varepsilon \bar{P}_0)$ are all nonnegative integers. But in this case we know (Theorem 1 of [5(b)] that there exists a finite-dimensional irreducible representation π of $\mathfrak{g}$ such that Λ is a weight of π while $\Lambda + \gamma$ is not a weight for any root $\gamma \varepsilon \mathfrak{F}^+(\bar{P}_0)$. Then it follows from Lemma 2 that if Λ' is any other weight of π, $\Lambda - \Lambda' \varepsilon \mathfrak{F}^+(\bar{P}_0)$. Therefore since $s\Lambda$ is a weight for every $s \varepsilon W$, we get the result.

Let us keep to the notation of the proof of Theorem 5 and put $\mathfrak{p}_- = \sum_{\gamma \varepsilon P_+} CX_{-\gamma}$. Since $[\mathfrak{p}_-, \mathfrak{p}_-] = \{0\}$ (Lemma 11) the subalgebra $\mathfrak{P}_-$ of $\mathfrak{B}$ generated by $(1, \mathfrak{p}_-)$ is abelian. We consider the tensor product $\mathfrak{P}_- \times \mathfrak{X}^*$.

LEMMA 21. *Let Γ denote the linear mapping $\mathfrak{P}_- \times \mathfrak{X}^*$ into $\mathfrak{B}^*$ such that $\Gamma(p \times x^*) = \pi(p)x^*$ $(p \varepsilon \mathfrak{P}_-, x^* \varepsilon \mathfrak{X}^*)$. Then Γ is a linear isomorphism of $\mathfrak{P}_- \times \mathfrak{X}^*$ onto $\mathfrak{B}^*$.*

We have seen during the proof of Theorem 4 that $\mathfrak{B}^* = \pi(\mathfrak{B}_+)1^*$ and $\mathfrak{X}^* = \pi(\mathfrak{X}_+)1^*$. Also we proved that $\mathfrak{M}_+ = \sum_{\gamma \varepsilon Q} \mathfrak{B}X_\gamma + \mathfrak{B}_+ Y_+$ is a maximal left ideal in $\mathfrak{B}_+$. Therefore $\mathfrak{B} = \mathfrak{B}_+ + \mathfrak{M}$, $\mathfrak{X} = \mathfrak{X}_+ + \mathfrak{Y}_{\Lambda_0}$ and $\mathfrak{M} \cap \mathfrak{B}_+ = \mathfrak{M}_+$. Put $\mathfrak{p}_+ = \sum_{\gamma \varepsilon P_+} CX_\gamma$ and let $\mathfrak{P}_+$ be the subalgebra of $\mathfrak{B}$ generated by $(1, \mathfrak{p}_+)$. Then if $\mathfrak{k}_+ = \mathfrak{k} \cap \mathfrak{g}_+$, $\mathfrak{g}_+ = \mathfrak{p}_+ + \mathfrak{p}_- + \mathfrak{k}_+$ and it follows from Lemma 12 of [5(c)] that $\mathfrak{B}_+ = \mathfrak{P}_- \mathfrak{P}_+ \mathfrak{X}_+$. But $[\mathfrak{p}_+, \mathfrak{k}] \subset \mathfrak{p}_+ \subset \mathfrak{M}_+$ and therefore

$$\mathfrak{P}_+ \mathfrak{X}_+ = \mathfrak{X}_+ + \mathfrak{P}_+ \mathfrak{p}_+ \mathfrak{X}_+ \subset \mathfrak{X}_+ + \mathfrak{M}_+.$$

Hence $\mathfrak{B}_+ = \mathfrak{P}_- \mathfrak{X}_+ + \mathfrak{M}_+$ and $\mathfrak{B}^* = \pi(\mathfrak{B}_+)1^* = \pi(\mathfrak{P}_-)\mathfrak{X}^* = \Gamma(\mathfrak{P}_- \times \mathfrak{X}^*)$. Moreover the mapping $(p \times p' \times x) \rightarrow pp'x$ $(p \varepsilon \mathfrak{P}_-, p' \varepsilon \mathfrak{P}_+, x \varepsilon \mathfrak{X}_+)$ defines a linear isomorphism μ of $\mathfrak{P}_- \times \mathfrak{P}_+ \times \mathfrak{X}_+$ onto $\mathfrak{B}_+$ [5(c), Lemma 12]. Therefore

$$\begin{aligned} \mathfrak{M}_+ &= \mathfrak{B}_+ \mathfrak{p}_+ + \mathfrak{B}_+ \mathfrak{Y}_+ = \mathfrak{P}_- (\mathfrak{P}_+ \mathfrak{p}_+) \mathfrak{X}_+ + \mathfrak{P}_- \mathfrak{P}_+ \mathfrak{Y}_+ \\ &= \mu((\mathfrak{P}_- \times (\mathfrak{P}_+ \mathfrak{p}_+) \times \mathfrak{X}_+) + (\mathfrak{P}_- \times \mathfrak{P}_+ \times \mathfrak{Y}_+)). \end{aligned}$$

Hence by considering $\mu^{-1}(\mathfrak{M}_+) \cap \mu^{-1}(\mathfrak{P}_- \mathfrak{X}_+)$ it is obvious that $\mathfrak{P}_- \mathfrak{X}_+ \cap \mathfrak{M}_+ = \mathfrak{P}_- \mathfrak{Y}_+$ and this shows that the kernel of Γ is zero. Therefore Γ is an isomorphism of $\mathfrak{P}_- \times \mathfrak{X}^*$ onto $\mathfrak{B}^*$.

For any linear function Λ on $\mathfrak{h}$ put $n(\Lambda) = \dim \mathfrak{B}_\Lambda^*$. We have seen above that $\mathfrak{X}^*$ is irreducible under $\pi(\mathfrak{k})$. Let $\Lambda_0, \Lambda_1, \cdots, \Lambda_r$ be all the distinct weights of the corresponding representation of $\mathfrak{k}$ and let μ_i denote the multiplicity of Λ_i so that $\mu_i = \dim(\mathfrak{B}_{\Lambda_i}^* \cap \mathfrak{X}^*)$ $(0 \leqq i \leqq r)$. Then we get the following result from the above lemma.

COROLLARY. *Let* $(\gamma_1,\cdots,\gamma_q)$ *be all the distinct totally positive roots of* $\mathfrak{g}$. *Then if* Λ *is any linear function on* $\mathfrak{h}$,

$$n(\Lambda)=\sum_{i=0}^{r}\mu_i n_i(\Lambda)$$

where $n_i(\Lambda)$ $(0\leqq i\leqq r)$ *is the number of distinct sequences* $(m\,,\cdots,m_q)$ *of nonnegative integers such that* $\Lambda=\Lambda_i-(m_1\gamma_1+\cdots+m_q\gamma_q)$.

Put $\mathfrak{X}_i^*=\mathfrak{B}_{\Lambda_i}^*\cap\mathfrak{X}^*$, $0\leqq i\leqq r$. Then $\mathfrak{X}^*$ is the direct sum of $\mathfrak{X}_i^*$, $0\leqq i\leqq r$ and we can choose a base $x_1^*,\cdots,x_d^*$ for $\mathfrak{X}^*$ which is composed of bases for the various $\mathfrak{X}_i^*$. Since $\mathfrak{P}_-$ is the universal enveloping algebra of the abelian Lie algebra $\mathfrak{p}_-$ and since $(X_{-\gamma_1},\cdots,X_{-\gamma_q})$ is a base for $\mathfrak{p}_-$, it follows that the elements $X_{-\gamma_1}^{m_1}X_{-\gamma_q}^{m_2}\cdots X_{-\gamma_q}^{m_q}$ $(m_1,\cdots,m_q\geqq 0)$ form a base for $\mathfrak{P}_-$. Hence it follows from the above lemma that $\pi(X_{-\gamma_1}^{m_1}\cdots X_{-\gamma_q}^{m_q})x_j^*$ $(m_1,\cdots,m_q\geqq 0, 1\leqq j\leqq d)$ is a base for $\mathfrak{B}^*$. Moreover it is obvious that if $x^*\,\varepsilon\,\mathfrak{X}_i^*$ $(0\leqq i\leqq r)$

$$\pi(X_{-\gamma_1}^{m_1}X_{-\gamma_2}^{m_2}\cdots X_{-\gamma_q}^{m_q})x^*\,\varepsilon\,\mathfrak{B}_\Lambda^*$$

where $\Lambda=\Lambda_i-(m_1\gamma_1+\cdots+m_q\gamma_q)$. Our assertion now follows immediately from these facts.

COLUMBIA UNIVERSITY,
NEW YORK, N. Y.

REFERENCES.

[1] V. Bargmann, "Irreducible unitary representations of the Lorentz group," *Annals of Mathematics*, vol. 48 (1947), pp. 568-640.

[2] G. Birkhoff and P. M. Whitman, "Representation of Jordan and Lie algebras," *Transactions of the American Mathematical Society*, vol. 65 (1949), pp. 116-136.

[3] E. Cartan, (a) "Les groupes projectifs qui ne lassent invariante aucune multiplicité plane," *Bulletin de la Société Mathématique de France*, vol. 41 (1913), pp. 53-96.

———, (b) "Sur certaines formes riemanniennes remarquables des géométries à groupe fondamental simple," *Annales Scientifiques de l'École Normale Supérieure*, vol. 44 (1927), pp. 345-467.

———, (c) "Groupes simples clos et ouverts et géométrie riemannienne," *Journal de Mathématiques Pures et Appliquées*, 9th Series, vol. 8 (1929), pp. 1-33.

———, (d) "Sur les domaines bornés homogènes de l'espace de n variables complexes," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität,* vol. 11 (1935), pp. 116-162.

———, (e) "La topologie des espaces représentatifs des groupes de Lie," *L'Enseignement mathématique,* vol. 35 (1936), pp. 177-200.

[4] I. M. Gelfand and M. I. Graev, (a) "Unitarnye predstavleniia veshchestvennoi unimoduliarnoi gruppy (osnovnye nevyrozhdennye serii)," *Izvestia Akademiia Nauk SSSR,* vol. 17 (1953), pp. 189-248.

———, (b) "Analog formuly Plansherelia dlia veshchestvennykh poluprostykh grupp Li," *Doklady Akademiia Nauk SSSR,* N. S., vol. 92 (1953), pp. 461-464.

[5] Harish-Chandra, (a) "On representations of Lie algebras," *Annals of Mathematics,* vol. 50 (1949), pp. 900-915.

———, (b) "On some applications of the universal enveloping algebra of a semisimple Lie algebra," *Transactions of the American Mathematical Society,* vol. 70 (1951), pp. 28-96.

———, (c) "Representations of a semisimple Lie group on a Banach space I," *ibid.,* vol. 75 (1953), pp. 185-243.

———, (d) "Representations of semisimple Lie groups. II," *ibid.,* vol. 76 (1954), pp. 26-65.

———, (e) "Plancherel formula for the 2×2 real unimodular group," *Proceedings of the National Academy of Sciences,* vol. 38 (1952), pp. 337-342.

———, (f) "Representations of semisimple Lie groups. V," *ibid.,* vol. 40 (1954), p. 1076.

———, (g) "Representations of semisimple Lie groups. VI," *ibid.,* vol. 40 (1954), p. 1078.

———, (h) "Integrable and square-integrable representations of a semisimple Lie group," *ibid.,* vol. 41 (1955), p. 314.

[6] G. W. Mackey, "Induced representations of locally compact groups. I," *Annals of Mathematics,* vol. 55 (1952), pp. 101-139.

[7] H. Weyl, *The Structure and Representation of Continuous Groups,* The Institute for Advanced Study, Princeton, 1935.

INVARIANTS OF FINITE GROUPS GENERATED BY REFLECTIONS.*

By Claude Chevalley.

---

1. An invertible linear transformation of a finite dimensional vector space V over a field K will be called a *reflection* if it is of order two and leaves a hyperplane pointwise fixed. A group G of linear transformations of V is a *finite reflection group* if it is a finite group generated by reflections. The operations of G extend to automorphisms of the symmetric algebra S of V by the rule $g(P)(x) = P(g^{-1}(x))$, $(P \varepsilon S, x \varepsilon V)$, and an element $P \varepsilon S$ such that $g(P) = P$ for all $g \varepsilon G$ is said to be *an invariant of* G. Our main purpose in this note is to prove the theorem:

(A) *Let G be a finite reflection group in a n-dimensional vector-space V over a field K of characteristic zero. Then the K-algebra J of invariants of G is generated by n algebraically independent homogeneous elements (and the unit).*

A vector space A is *graded* by subspaces A^i, (i positive integer), if it is the direct sum of the A^i. The degree d^0P of $P \varepsilon A$ is the smallest integer j such that $P \varepsilon \sum_{i \leqq j} A^i$; the elements of A^i are the homogeneous elements of degree i. When the A^i are finite dimensional, the Poincaré series of A in the indeterminate t is defined as

$$P_t(A) = \sum_{i \geqq 0} \dim . A^i \cdot t^i.$$

In particular S is graded in the obvious way and $P_t(S) = (1 - t)^{-n}$. Let F be the ideal generated by the homogeneous elements of strictly positive degrees in J. Then the grading of S induces a grading of the quotient space S/F. Since F is invariant under G, the operations of G in S induce automorphisms of S/F. We shall also prove:

(B) *Let $I_1, \cdots, I_n$ be a minimal system of homogeneous generators of J and let m_i be the degree of I_i, $(1 \leqq i \leqq n)$. Then*

$$P_t(S/F) = (1 - t)^{-n} \cdot \prod_{i=1}^{i=n} (1 - t^{m_i}).$$

---

* Received June 9, 1955.

The product of the m_i *is equal to the order of* G *and to the dimension of* S/F. *The natural representation of* G *in* S/F *is equivalent to the regular representation.*

2. Two lemmas. In this paragraph, the characteristic p of the infinite groundfield K is allowed to be $\neq 0$ and G denotes a finite reflection group in V whose order N is prime to p.[1] To any element $P \varepsilon S$ we can then associate its average over G:

$$M(P) = 1/N \sum_{g \varepsilon G} g(P).$$

Lemma 1. *Let* $U_1, \cdots, U_m$ *be invariants of* G *such that* U_1 *does not belong to the ideal generated in* J *by* $U_2, \cdots, U_m$. *Let* P_i, $(1 \leqq i \leqq m)$, *be homogeneous elements of* S *satisfying a relation* $\sum_1^m P_i \cdot U_i = 0$. *Then* $P_1 \varepsilon F$.

If $d^0 P_1 = 0$, then it follows from the assumption and from the relation

$$M(P_1) \cdot U_1 + \cdots + M(P_m) \cdot U_m = 0$$

that $P_1 = M(P_1) = 0$. Assume now $d^0 P_1 > 0$ and the lemma to be true for all relations $\sum_1^m Q_i \cdot U_i = 0$ with homogeneous Q_i and $d^0 Q_1 < d^0 P_1$. Let s be a reflection of G leaving pointwise fixed a hyperplane with equation $L = 0$. Then $s(P_i) - P_i = L \cdot Q_i$, $(Q_i \varepsilon S,\ i = 1, \cdots, m)$, and

$$Q_1 \cdot U_1 + \cdots + Q_m \cdot U_m = 0$$

whence, by induction, $Q_1 \varepsilon F$ or, otherwise said, $s(P_1) \equiv P_1 \operatorname{mod.} F$; the group G being generated by reflections, we have then $g(P_1) \equiv P_1 \operatorname{mod.} F$ for any $g \varepsilon G$, whence $P_1 \equiv M(P_1) \operatorname{mod.} F$; since P_1 is homogeneous of strictly positive degree, the same is true for $M(P_1)$; therefore $M(P_1) \varepsilon F$ and $P_1 \varepsilon F$.

Lemma 2. *Assume* K *to be a perfect field. Let* I_i, $(1 \leqq i \leqq m)$, *be homogeneous invariants which form an ideal basis of* F,[2] *with* $m_i = d^0 I_i$ *prime to* p *for* $i \leqq r$. *Then* $I_1, \cdots, I_r$ *are algebraically independent.*

Let us suppose the lemma to be false and let $H(I_1, \cdots, I_r) = 0$ be a non trivial relation of minimal degree between $I_1, \cdots, I_r$ where $H(y_1, \cdots, y_r)$

---

[1] In this paper, we are primarily interested in the case $p = 0$, but Lemma 2 will be used in a forthcoming paper of A. Borel, to appear in Jour. Math. Pur. Appl.

[2] This always exists since by the classical theorem for invariants of a finite group, J is a finitely generated K-algebra.

is a polynomial in r letters y_i. We may assume that there exists an integer h such that for any monomial $y_1^{k_1} \cdot \cdots \cdot y_r^{k_r}$ of H we have

$$k_1 \cdot m_1 + \cdots + k_r \cdot m_r = h.$$

The partial derivatives $\partial H/\partial y_i$ are not all zero, because otherwise (for $p \neq 0$, the only case for which it is not obvious), K being perfect, H would be the p-th power of a polynomial H^*, and $H^*(I_1, \cdots, I_r) = 0$ would be a non trivial relation of strictly smaller degree. Set

$$H_i = \partial H/\partial y_i \, (I_1, \cdots, I_r), \qquad (1 \leqq i \leqq r);$$

then $H_1, \cdots, H_r$ are in J and not all zero; after a possible permutation of indices, we may assume that they belong to the ideal generated in J by the first s of them, but that none of $H_1, \cdots, H_s$ belongs to the ideal generated by the other ones in J. Set

$$H_{s+j} = \sum_{j=1}^{i=s} V_{j,i} H_i.$$

Let x_k, $(1 \leqq k \leqq n)$, be coordinates in V. Since

$$\sum_{i=1}^{i=r} H_i \cdot (\partial I_i/\partial x_k) = 0, \qquad (1 \leqq k \leqq n),$$

we have by Lemma 1

$$\partial I_i/\partial x_k + \sum_{j=1}^{j=r-s} V_{j,i}(\partial I_{s+j}/\partial x_k) \in F, \qquad (1 \leqq i \leqq s; 1 \leqq k \leqq n)$$

(the left hand sides are homogeneous in the x_k by the above remark on the monomials of H). Multiplying this relation by x_k and adding the relations thus obtained, we get

$$m_i I_i + \sum_{j=1}^{j=r-s} V_{j,i} m_{s+j} I_{s+j} = \sum_{l=1}^{l=m} A_{i,l} I_l, \qquad (1 \leqq i \leqq s).$$

where the $A_{i,l}$ are forms belonging to the ideal generated by $x_1, \cdots, x_n$. For reasons of homogeneity, we have $A_{i,l} = 0$ if I_l is not of strictly lower degree than I_i; m_i being prime to p for $i \leqq r$, we see that I_i belongs to the ideal generated by the other I_j, which is a contradiction. Thus $I_1, \cdots, I_r$ are algebraically independent.

3. Proofs of Theorems (A) and (B). We assume again the ground-field to be of characteristic zero and denote as in Lemma 2 by $I_1, \cdots, I_m$ homogeneous invariants of G forming an ideal basis of F. By Lemma 2 they are algebraically independent, whence also $m \leqq n$. Using averages over G, it is readily seen by induction on the degree that the unit and the I_i generate J and thus, to finish the proof of (A), there remains to show that $m \geqq n$.

Let $x_1, \cdots, x_n$ be coordinates in V and let $K(x)$ be the field of rational functions in the x_i. It is acted upon in a natural way by G and we denote by L the subfield of elements invariant under G. Then $K(x)$ is a Galois extension of L, with Galois group G and L has also transcendence degree n over K. On the other hand, G being *finite*, every invariant in $K(x)$ is classically the quotient of two invariant polynomials; thus $L = K(J)$ is generated by the I_i, and $m \geqq n$.

LEMMA 3. *Let $P_1, \cdots, P_s$ be homogeneous elements of S whose residue classes* mod F *are linearly independent over K in S/F. Then $P_1, \cdots, P_s$ are linearly independent over $K(J)$.*

Let $V_1 \cdot P_1 + \cdots + V_s \cdot P_s = 0$ be a relation with $V_i \varepsilon K(J)$, $(1 \leqq i \leqq s)$. We have to prove that $V_i = 0$ for all i and it is enough to consider the case where the V_i are homogeneous elements of J such that $d^0V_i + d^0P_i$ is equal to a constant h independent of i.

By the degree of the monomial $I_1^{k_1} \cdot \cdots \cdot I_n^{k_n}$ we mean its degree as element of S, i.e. $k_1m_1 + \cdots + k_nm_n$. Let S_j, $(j = 1, 2, \cdots)$, be the different monomials in the I_i arranged by increasing degrees, with $S_1 = 1$. We have

$$V_i = \sum_{j \geqq 0} k_{ij} S_j, \qquad (k_{ij} \varepsilon K, k_{ij} = 0 \text{ for } d^0V_i \neq d^0S_j, \ (1 \leqq i \leqq n)),$$

and our relation may be written

$$\sum_{j \geqq 0} W_j \cdot S_j = 0, \qquad (W_j = \sum_{i=1}^{i=s} k_{ij} P_i),$$

where W_j is homogeneous, of degree equal to $h - d^0S_j$. Assume that $k_{ij} = 0$ for $1 \leqq i \leqq s$ and $j < t$. Since by Theorem A the monomial S_t does not belong to the ideal generated in J by the S_j with $j > t$, we have by Lemma 1 $W_t \varepsilon F$ and the hypothesis gives then $k_{it} = 0$ for $i = 1, \cdots, s$. This proves by induction on j that $k_{ij} = 0$ for all i, j, and the lemma.

We now come to the proof of (B). The field $K(x)$ being a normal extension of $K(J)$ with Galois group G, has degree N over $K(J)$, hence the dimension of S/F over K is finite. Let $A_1, \cdots, A_q$ be homogeneous polynomials whose residue classes mod F form a basis of S/F. By induction on the degree we see that every $P \varepsilon S$ may be expressed as linear combination of the A_i with coefficients in J, and this expression is unique in view of Lemma 3. Hence

$$P_t(S)=P_t(S/F)\cdot P_t(J);$$

but $P_t(S)=(1-t)^{-n}$ and Theorem A gives $P_t(J)=\prod_1^n(1-t^{m_i})^{-1}$, whence the first assertion of (B). We may also write

$$P_t(S/F)=\prod_{i=1}^{i=n}(1+t+t^2+\cdots+t^{m_i-1})$$

and, setting $t=1$, we get $\dim. S/F=m_1\cdot\cdots\cdot m_n$. Since every element of $K(x)$ may be written as the quotient of a polynomial by an invariant polynomial, it also follows from the above and Lemma 3 that the A_i form a basis of $K(x)$ over $K(J)$, whence $N=\dim. S/F$.

We have for $g\,\varepsilon\, G$

$$g(A_i)=\sum_{j=1}^{j=N}a_{ij}(g)A_j, \qquad (i=1,\cdots,N),$$

where the $a_{ij}(g)$ are homogeneous elements of J and where $a_{ii}(g)\,\varepsilon\, K$ by homogeneity. The matrices $(a_{ij}(g))$ describe the natural representation of G in $K(x)$, considered as vector space over $K(J)$. If we reduce the coefficients $\operatorname{mod} F$ we get the natural representation of G in S/F, considered as vector space over K; this reduction does not affect the diagonal coefficients, hence both representations have the same character and are equivalent. But G is the Galois group of the normal extension $K(x)$ of $K(J)$, so that the former representation is equivalent to the regular representation, which proves the last statement of (B).

COLUMBIA UNIVERSITY.

ON THE ARITHMETIC STRUCTURE OF A CLASS OF COMMUTATIVE SEMI-GROUPS.*

By Lee M. Sonneborn.

1. Introduction. Let S be a commutative semi-group [Jacobson [3], Chapter IV] in which both chain conditions hold for non-associated elements, and consider a multiplicative diophantine system [Skolem [6], Chapter IV] of the form

$$x_1^{a_{i1}}x_2^{a_{i2}}\cdots x_k^{a_{ik}} = y_1^{b_{i1}}y_2^{b_{i2}}\cdots y_h^{b_{ih}} \qquad (i = 1, 2, \cdots, m) \tag{1}$$

where the a_{ij} and b_{ij} are non-negative integers and the x_i and y_i are indeterminates over S.

It was shown by E. T. Bell [1] that if S is the semi-group of positive integers, the solution of (1) may be reduced to the special case

$$x_1x_2 = y_1y_2, \qquad \text{for which} \tag{2}$$

$$x_1 = u_1v_1, \quad x_2 = u_2v_2, \quad y_1 = u_1u_2, \quad y_2 = v_1v_2. \tag{3}$$

We investigate here what restrictions are imposed on S if, whenever (2) holds in S, elements u_1, u_2, v_1, v_2 exist in S such that (3) is satisfied. Such a system S is said by Birkhoff [2] to satisfy the Riesz Refinement Postulate after F. Riesz [5]. We shall prove that upon identifying associated elements, S becomes the sub-direct product of a distributive lattice. and a Gaussian semi-group [Jacobson, loc. cit.].

2. We say a "divides, or does not divide," b ($a|b$ or $a\nmid b$) if there exists, or does not exist an element, c, of S such that $ac = b$. If a divides b and b divides a, then a and b are called associates ($a \sim b$). Otherwise $a \not\sim b$. If S contains an element e_a such that $e_a a = a$, then e_a is called an identity for a. If $e|x$ for all x in S, then e is called a unit. It is not, in general, true that two associate elements differ by a unit, as is shown by the system

$$\begin{array}{c|ccc} & 1 & a & b \\ \hline 1 & 1 & a & b \\ a & a & a & b \\ b & b & b & a \end{array}$$

in which $a \sim b$, but 1 is the only unit.

---

* Received May 18, 1955.

We consider now an associative, commutative semi-group S in which

(1) *Every ascending or descending chain of non-associated divisors of any element, a, of S is of finite length.*

(2) $ab = cd$ *if and only if there exist elements* x, y, z, w *in* S *such that* $a = xy$, $b = zw$, $c = xz$, *and* $d = yw$.

On applying (2) to the equation $aa = aa$, we find that

$$a = x_1 y_1 = z_1 w_1 = x_1 z_1 = y_1 w_1.$$

If $x_1 \not\sim a$, then $x_1 x_1 = x_1 x_1$ implies

$$x_1 = x_2 y_2 = z_2 w_2 = x_2 z_2 = y_2 w_2.$$

Continuing in this fashion we obtain a chain of divisors of $a, x_1, x_2 \cdots x_i \cdots$, in which each element divides all the preceding elements. By condition (1) above, only a finite number of the x_i can be non-associates. Hence $x_n \sim x_{n'}$ for some n and n'. But then $a = kx_n = lx_{n'}$ for some k and l, and $x_n = mx_{n'}$ and $x_{n'} = px_n$. Hence $mpa = kmpx_n = kmx_{n'} = kx_n = a$, so mp is an identity for a. But a was arbitrary; hence we have proved

(3) *For each element a of S, there exists an e_a, called an identity for a, such that* $e_a a = a$.

Condition (3) implies that the relation $\sim$ is reflexive. *Hence $\sim$ is an equivalence relation*; for its symmetry and transitivity are obvious from the definition.

THEOREM 1. *The system S', whose elements are the equivalence classes of S under $\sim$, is an associative commutative semi-group in which*

(1′) *Any ascending or descending chain of proper divisors of an element of S' is of finite length*

(2′) $ab \sim cd$ *if and only if* $a \sim xy$, $xz \sim st \mid c$, $c \sim su$, *and*

$$b \sim zw,\quad yw \sim uv \mid d,\quad d \sim tv.$$

(3′) *For any a in S', there is an e_a in S' such that* $e_a a = a$.

Proof. Since $\sim$ is defined by division, $a \sim a'$ and $b \sim b'$ imply $a = ka'$, $a' = la$, $b = mb'$, $b' = nb$ for some k, l, m, and n, and hence $ab = kma'b'$ and $a'b' = ln\,ab$ which gives $ab \sim a'b'$. Thus multiplication is well defined in S'. Properties (1′) and (3′) then follow immediately from conditions (1) and (3) in S. The proof of (2′) rests on two lemmas.

Consider the co-ideal $\{a\}$, [MacKenzie [4]], consisting of all divisors of the element a. Clearly $\{a\} \subseteq \{b\}$ if and only if $a|b$, and $\{a\} = \{b\}$ if and only if $a \sim b$.

LEMMA 1. *$\{a\}\{b\} = \{ab\}$ where $\{a\}\{b\}$ is the set of all elements a_1b_1 with $a_1 \varepsilon \{a\}$, and $b_1 \varepsilon \{b\}$.*

Proof. Trivially $\{a\}\{b\} \subseteq \{ab\}$. Let $c \varepsilon \{ab\}$. Then $c|ab$ and therefore $ab = cd$ for some d. But then, by (1) $a = xy$, $b = zw$, $c = xz$, $d = yw$, where c has been displayed as xz and $x|a$ and $z|b$. Let $a_1 = x$, $b_1 = z$. Thus $c \varepsilon \{a\}\{b\}$.

The following two corollaries, amounting to little more than restatement, will be useful later:

COROLLARY 1. *If $\{c\} \subseteq \{ab\}$, then $c = c_1c_2$ where $c_1 \varepsilon \{a\}$, and $c_2 \varepsilon \{b\}$.*

COROLLARY 2. *If $\{cd\} = \{ab\}$, then $ab = a_1b_1$ where $a_1 \varepsilon \{c\}$ and $b_1 \varepsilon \{d\}$.*

LEMMA 2. *If $\{ab\} = \{cd\}$, then there exist a', b', c', and d' such that $ab = a'b'$, $cd = c'd'$, and $\{a'\} = \{c'\} \subseteq \{c\}$ and $\{b'\} = \{d'\} \subseteq \{d\}$.*

Proof. We form a chain which is forced to terminate by condition (1) in S. By Corollary 2, we choose a_1 and b_1 such that $a_1b_1 = ab$ and $\{a_1\} \subseteq \{c\}$ and $\{b_1\} \subseteq \{d\}$. Then from $\{a_1b_1\} = \{cd\}$ we choose c_1 and d_1 so that $c_1d_1 = cd$ and $\{c_1\} \subseteq \{a_1\} \subseteq \{c\}$ and $\{d_1\} \subseteq \{b_1\} \subseteq \{d\}$. Continuing in this fashion, we obtain a_2 and b_2 such that $ab = a_1b_1 = a_2b_2$ and

$$\{a_2\} \subseteq \{c_1\} \subseteq \{a_1\} \subseteq \{c\}, \text{ and } \{b_2\} \subseteq \{d_1\} \subseteq \{b_1\} \subseteq \{d\}$$

and, eventually,

$$ab = a_1b_1 = a_2b_2 = a_3b_3 = \cdots = a_nb_n = \cdots$$

and

$$cd = c_1d_1 = c_2d_2 = c_3d_3 = \cdots = c_nd_n = \cdots,$$

in which

$$\{a_n\} \subseteq \{c_{n-1}\} \subseteq \{a_{n-1}\} \subseteq \{c_{n-2}\} \subseteq \cdots \subseteq \{c\},$$

and

$$\{b_n\} \subseteq \{d_{n-1}\} \subseteq \{b_{n-1}\} \subseteq \{d_{n-2}\} \subseteq \cdots \subseteq \{d\}.$$

But now, by virtue of condition (1) in S, there is an m such that $\{a_m\} = \{c_m\}$ and $\{b_m\} = \{d_m\}$. Let $a_m = a'$, $b_m = b'$, $c_m = c'$ and $d_m = d'$ and the lemma is proved.

We may now conclude the proof of (2′) and hence of Theorem 1 as follows: We have $\{ab\} = \{cd\}$ if and only if $ab \sim cd$. Then by Lemma 2,

$ab \sim cd$ implies $ab = a'b'$, and $c'd' = cd$ where $a' \sim c'$, $b' \sim d'$ and $a' | c$, and $b' | d$. But by condition (2) in S, we have

$$a = xy,\ a' = xz \sim st = c' | c,\ c = su;$$

$$b = zw,\ b' = yw \sim uv = d' | d,\ d = tv.$$

Clearly equality implies equivalence, so all equalities in the above may be replaced by equivalences, thus completing the proof of Theorem 1.

3. In the remainder of the paper the system S' is considered where equivalence in S is replaced by equality in S'. S' is then an associative, commutative semi-group in which

(1′) *Any ascending or descending chain of proper divisors of any element of S' is of finite length.*

(2′) $ab = cd$ *if and only if there exist* x, y, z, w, s, t, u, *and* v *such that*

$$a = xy,\ xz = st | c,\ c = su;$$

$$b = zw,\ yw = uv | d,\ d = tv.$$

(3′) *For any* a *in* S', *there is an identity for* a, *and*

(4′) *If* $a | b$ *and* $b | a$, *then* $a = b$.

THEOREM 2. *There is an element* 1 *in* S' *such that* $1a = a$ *for all* a *in* S.

Proof. Clearly any divisor of an identity for a is an identity for a, by virtue of condition (4′). Also any two identities for the same element have a common divisor, for if $e_a a = e_a' a = a$, we have from (2′) $e_a = xy$, $xz = st | e_a'$, $e_a' = su$ etc., where $x | e_a$ and $x | e_a'$. Now, if $e_a a = a$ and $e_b b = b$, e_a and e_b have a common divisor which is an identity for a and b, since $e_a ab = e_b ab = ab$. Consider now a maximal chain of divisors of e_a where $e_a a = a$. It has a smallest element e_{0_a} which has the property that it divides all elements of the chain. If a second chain is picked, its smallest element e_{0_a}' must equal e_{0_a} since the two minimal elements must have a common divisor. Hence e_{0_a} depends only on a, not on the chain selected. For any other element b there is similarly an identity e_{0_b}. But then, by the above e_{0_a} and e_{0_b} have a common divisor which must be equal to both by minimality. Hence $e_{0_a} x = x$ for all x and e_{0_a} is the element 1 with properties stated in Theorem 2.

4. We define an element p, to be irreducible if the only divisors of p are identities for p and p itself. Clearly any element of S' can be written as a finite product of irreducibles in at least one way because of the descending chain condition.[7]

THEOREM 3. *The set of idempotent elements of S' is identical with the set of elements which are identities.*

Proof (a) *Any non-idempotent irreducible has all its powers distinct.* For suppose p is irreducible and $p \neq p^2$ and $p^n = p^{n+m}$ for some n and m. Then $p^n \mid p^{n+1} \mid p^{m+n} = p^n$ implies by (4′) that $p^n = p^{n+1}$. We pick the smallest n for which this is true ($n \geqq 2$) and eventually arrive at the contradiction that $p^n = p^{n+1}$ implies $p^{n-1} = p^n$. Now $p^{n+1} = p^n$ is the same as $p \cdot p^n = p \cdot p^{n-1}$ where the only divisors of p are p and identities for p. But

$$p = xy,\ xz = st \mid p,\ su = p;$$

$$p^n = zw,\ yw = uv \mid p^{n-1},\ tv = p^{n-1}.$$

Clearly, x or y, and s or u, must be p since p is non-idempotent and irreducible.

Case I. $x = s = p$; $y = e_p'$, $u = e_p$, where e_p and e_p' are identities for p. Then z and t must be identities for p, say e_p'' and e_p'''. Then $p^n = zw = e_p''w$ and $e_p'w = yw \mid p^{n-1}$. But then $e_p''e_p'w \mid e_p''p^{n-1} = p^{n-1}$ and $e_p'e_p''w = e_p'p^n = p^n$ from which $p^n \mid p^{n-1}$, or $p^n = p^{n-1}$, the desired contradiction.

Case II. $y = s = p$; $x = e_p$, $u = e_p'$, $t = e_p''$, $z = p$. But then, similarly $p^n = zw = pw$ and $yw = pw \mid p^{n-1}$.

Case III. $x = u = p$; $y = e_p$, $z = e_p'$, $s = e_p''$ and $t = p$. Then $p^n = zw = e_p'w$, $yw = e_pw \mid p^{n-1}$. But $e_pp^n = p^n = e_pe_p'w$ and

$$e_p'yw = e_p'e_pw \mid e_p'p^{n-1} = p^{n-1} \text{ or } p^n = e_pe_p'w \mid p^{n-1}.$$

Case IV. $y = u = p$; $x = e_p$, $s = e_p'$. Two cases are possible. First if $z = p$, then $p^n = pw$, $yw = pw \mid p^{n-1}$, or $p^n = p^{n-1}$. Otherwise $z = e_p''$ in which case $p^n = zw = e_p''w$ and $pw \mid p^{n-1}$ whence $p \cdot p^n = p \cdot e_p''w = pw \mid p^{n-1}$ from which $p^{n+1} \mid p^{n-1}$ which is false and completes the proof of (a).

(b) *Any divisor of an idempotent is idempotent.* For let a be idempotent and $a = pb$ where p is irreducible. Then $a^n = p^nb^n = a$. Hence $p^n \mid a$ for all n. But the p^n are an infinite ascending chain of divisors of a unless p is idempotent. Hence p is idempotent. But then any irreducible divisor of an idempotent is idempotent. But any divisor of an idempotent is a product of irreducibles and is hence idempotent.

(c) *Any identity is idempotent.* Consider an element e_a with the property that $e_a a = a$. Let p, an irreducible, divide e_a. Then $p^n \mid a$ for all a and hence p is idempotent. But e_a is the product of a finite number of such idempotents, and is thus idempotent. But clearly any idempotent is an identity (for itself) and so the set of identities is the same as the set of idempotents completing the proof of Theorem 3.

THEOREM 4. *The identities of S' form a distributive lattice.*

Proof. Let I be the set of idempotents partially ordered by $x \leqq y$ if $y \mid x$, and consider the product ab with a and b in I. Then ab is in I and $ab \leqq a$, $ab \leqq b$. Assume x in I and $x \leqq a$, $x \leqq b$. Then $x = ka$ and $ax = aak = ak = x$ and $abx = x$. Hence $ab \mid x$ and $x \leqq ab$. Thus $ab = a \cap b$, and I is a lattice, for it is a partially ordered set with a unit element 1, the descending chain condition, and is closed under finite intersection. I is distributive for if $a \geqq b \cap c$, there exist b_1 and c_1 such that $a = b_1 \cap c_1$ with $b_1 \geqq b$ and $c_1 \geqq c$. This is easily shown to be equivalent to distributivity in a lattice, and, rewritten in terms of division, is precisely Corollary 1 of Section 2. Now, $ab = cd$ if and only if $a = xy$, $b = zw$, $c = xz$, $d = yw$, for distributivity in a lattice is also equivalent to $a \cap b = c \cap d$ if and only if $a = (a \cup c) \cap (a \cup d)$, $b = (b \cup c) \cap (b \cup d)$, $c = (a \cup c) \cap (b \cup c)$ and $d = (a \cup d) \cap (b \cup d)$. Hence *condition* (2) *of S holds among the identities of S'.*

5. THEOREM 5. *If an element of S' is factored into irreducibles in two ways, the non-idempotents and their powers are identical in the two factorizations.*

Proof. Let p, p_1, p_2 denote non-idempotent irreducibles.

(a) If $p \mid ab$ then $p \mid a$ or $p \mid b$. For let $pk = ab$. Then by condition (2′) on S',

$$\begin{array}{ll} p = xy & xz = st \mid a, \\ k = zw & yw = uv \mid b, \end{array} \qquad \text{where } x \text{ or } y = p \text{ and } x \mid a, y \mid b.$$

(b) Any non-idempotent irreducible has no non-idempotent divisors except itself.

(c) In two factorizations of an element into irreducibles the distinct non-idempotents occurring are the same. For, let p occur in one factorization, then by (a) it must divide a p' in the second factorization and by (b) it must be the same element.

(d) No non-idempotent irreducible divides any power of any other non-idempotent. For let $p_1 \neq p_2$ and $p_1 | p_2^{\alpha}$ where α is minimal with this property. Then $p_1 | p_2^{\alpha-1} p_2$. But by (a), $p_1 | p_2$ or $p_1 | p_2^{\alpha-1}$ both of which are false.

(e) If $p^n | ab$ and $p \nmid a$ then $p^n | b$. We prove this by induction on n. By (a) this is true for $n = 1$. Let m be the smallest integer for which $p^m | ab$, $p \nmid a$ and $p^m \nmid b$. By the induction hypothesis, $p^{m-1} | ab$ and $p \nmid a$ implies $p^{m-1} | b$. From $kp^m = ab$, we have by (2′)

$$p^m = xy, xz = st \,|\, a, su = a;$$

$$k = zw, yw = uv \,|\, b, tv = b.$$

But $p \nmid a$ and hence x, z, s, t and u are not divisible by p. But then $p^m = xy$ where $p \nmid x$ and hence $p^{m-1} | y$. Hence $p \cdot p^{m-1} = xy$ from which

$$p^{m-1} = x'y', x'z' = s't' \,|\, x, s'u' = x;$$

$$p = z'w', y'w' = u'v' \,|\, y, t'v' = y.$$

But p cannot divide x, or hence x', z', t' or u'. But if $p \nmid x'$, then $p^{m-1} | y'$ and if $p \nmid z'$, $p = w'$. Thus $p \cdot p^{m-1} = p^m | w'y' | y$ and $y | b$, so $p^m | b$ contrary to assumption. Now, if an element a of S' has two factorizations

$$a = e_\alpha p_1^{\alpha_1} \cdots p_r^{\alpha_r} = e_\beta p_1^{\beta_1} \cdots p_r^{\beta_r}$$

in which e_α and e_β are the idempotent parts, $\alpha_i = \beta_i$ for all i since $p_i^{\alpha_i} | kp_i^{\beta_i}$ where $p_i \nmid k$. This completes the proof of Theorem 5.

The set of idempotents which divide a, an element of S', has a maximal element e_a. Say it has a second e_a'. Then $e_a a = e_a' a = a = e_a e_a' a$. But since e_a and e_a' are maximal they are equal. Hence we may factor the element a uniquely into e_a, its maximal idempotent divisor, and powers of irreducible non-idempotents. Let $a = e_a p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, and $b = e_b p_1^{\beta_1} \cdots p_r^{\beta_r}$ (α_i and $\beta_i \geqq 0$) be factorizations in this unique fashion. Then

$$ab = e_a e_b p_1^{\alpha_1+\beta_1} \cdots p_r^{\alpha_r+\beta_r}$$

is the factorization of their product. For clearly $e_a e_b | ab$, and $e_{ab} | ab$ implies, by Corollary 1 of Section 2, that $e_{ab} = e_a' e_b'$ such that $e_a' | e_a$ and $e_b' | e_b$. But this in turn implies that $e_{ab} = e_a' e_b' | e_a e_b | e_{ab}$. Hence $e_a e_b = e_{ab}$.

THEOREM 6. *$ab = cd$ if and only if there exist x, y, z, and w such that $a = xy$, $b = zw$, $c = xz$ and $d = yw$.*

Proof. $a = e_a p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, $b = e_b p_1^{\beta_1} \cdots p_r^{\beta_r}$, $c = e_c p_1^{\gamma_1} \cdots p_r^{\gamma_r}$ and

$d = e_d p_1^{\delta_1} \cdots p_r^{\delta_r}$. By the above $ab = cd$ if and only if $e_a e_b = e_c e_d$ and $\alpha_i + \beta_i = \gamma_i + \delta_i$. From the remark on distributive lattices $e_a = x_1 y_1$, $e_b = z_1 w_1$, $e_c = x_1 z_1$ and $e_d = y_1 w_1$, and from the truth in integers (or arbitrary Gaussian semi-groups), the truth for irreducibles is immediate. The product of the separate parts clearly suffices.

THEOREM 7. *The system S' is the sub-direct product of a distributive lattice and a Gaussian semi-group.*

Proof. Replace $a = e_a p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ by the pair $(e_a, p_1^{\alpha_1} \cdots p_r^{\alpha_r})$. Then $ab = (e_a \cap e_b,\ p_1^{\alpha_1+\beta_1} \cdots p_r^{\alpha_r+\beta_r})$. However, the product is not necessarily direct. For consider the set $R = \{(e_0, 1), (e_1, 1), (e_1, p^n)\ (n = 1, 2, 3, \cdots)\}$ with $e_0 \cap e_1 = e_1$. All the postulates are satisfied, but R is not a direct product, for the $\{(e_0, p^n)\}$ are missing.

The question of whether or not any diophantine system may be reduced as in Bell [1] to the special case considered here will be treated elsewhere.

The author wishes to thank Professor Morgan Ward who suggested the problem, and whose helpful suggestions and encouragement are greatly appreciated.

CALIFORNIA INSTITUTE OF TECHNOLOGY.

REFERENCES.

[1] E. T. Bell, "Reciprocal arrays and diophantine analysis," *American Journal of Mathematics*, vol. 55 (1933), pp. 50-66.

[2] G. Birkhoff, "Lattice-ordered groups," *Annals of Mathematics*, vol. 43 (1942), p. 328.

[3] N. Jacobson, "Lectures in Abstract Algebra," vol. 1, Chap. IV, Van Nostrand, New York, 1951.

[4] R. E. MacKenzie, "Commutative semi-groups," *Duke Mathematical Journal*, vol. 21 (1954), pp. 471-477.

[5] F. Riesz, "Sur quelques notions fondamentales dans la théorie générale des operations lineaires," *Annals of Mathematics*, vol. 41 (1940), p. 175.

[6] Th. Skolem, "Diophantische Gleichungen," *Ergebnisse der Mathematik und ihrer Grenzgebiete*, vol. 5 (1938), Chap. IV.

ON A PROBLEM OF POINCARÉ CONCERNING RICCATI'S EQUATION.*

By PHILIP HARTMAN and AUREL WINTNER.

---

1. For large positive t, consider the differential equation

$$x'' + f(t)x = 0 \tag{1}$$

in which $f(t)$ is a real-valued, continuous function having the property that

$$\omega = \lim_{t \to \infty} f(t) \text{ exists} \tag{2}$$

(as a finite limit). Then equation (1) can be considered to be a "small" perturbation of the equation with constant coefficients

$$x'' + \omega x = 0. \tag{3}$$

This prototype of (1) has the pair of solutions $x = e^{\pm\lambda t}$, $x = e^{\pm i\lambda t}$ or $x = 1, t$ according as $\omega = -\lambda^2 < 0$, $\omega = \lambda^2 > 0$ or $\omega = 0$. Correspondingly, an assertion of Poincaré [9], pp. 227-236, implies that (1), in the case (2), has solutions $x = z(t)$ with a logarithmic derivative

$$w = z'/z \tag{4}$$

satisfying

$$w^2(t) \to -\omega \text{ as } t \to \infty. \tag{5}$$

But the indications given by Poincaré are incomplete, and the purpose of this paper is to examine the validity of the assertion (when $\omega = f(\infty) \geqq 0$).

What is at stake is the existence, and not the evaluation, of a limit, $w(\infty)$, of (4) as $t \to \infty$. For if it is assumed that $w(\infty)$ exists, then $w^2(\infty) = -\omega$ (so that (5) holds). In order to see this, note that (4) satisfies the Riccati equation

$$w' + w^2 + f(t) = 0$$

(at those t-values at which $z \neq 0$). Hence the assumptions that $f(\infty) = \omega$ and $w(\infty)$ exist imply that $w'(t)$ tends to $-w^2(\infty) - \omega$ as $t \to \infty$. But the existence of $w(\infty)$ means that $w'(t)$ is improperly integrable at $t = \infty$, and so the limit of $w'(t)$ at $t = \infty$ must be 0, which proves that $w^2(\infty) = -\omega$.

---

* Received May 5, 1955.

Poincaré's statement is correct if $f(\infty) < 0$ and, in this case,* every non-trivial ($\not\equiv 0$), real- or complex-valued solution $x = z(t)$ of (1) satisfies (4)-(5); see Perron [6]. When

$$(6) \qquad \omega = 0,$$

then (1) can be non-oscillatory or oscillatory, where, in the terminology of A. Kneser, (1) is called non-oscillatory if one (hence, by Sturm's separation theorem, every) real-valued solution ($\not\equiv 0$) of (1) does not vanish for large t. In the non-oscillatory case (2), (6) of (1), it is known that every (real- or complex-valued) solution $x = z(t)$ ($\not\equiv 0$) of (1) satisfies (4), (5); see [2], p. 391, Theorem B.

If (1) is oscillatory and $x = z(t)$ is a real-valued solution, then the ratio (4) is undefined at the zeros of $z(t)$. On the other hand, if $x = x(t)$ and $x = y(t)$ are two real-valued, linearly independent solutions of (1) (which, therefore, do not vanish simultaneously), then the complex-valued solution

$$(7) \qquad z = x(t) + iy(t)$$

of (1) has no zeros (so that (4) is defined for all t). Perron [7] has shown that, nevertheless, Poincaré's assertion in the case (6) is false; that is, that there exist differential equations (1) which have coefficient functions satisfying (2), (6) (and which are necessarily oscillatory) but are such that no (complex-valued) solution $x = z(t)$ satisfies (4), (5). Strictly speaking, Poincaré's paper assumes that $f(t)$ is a rational function, but this is of no avail (in this connection, cf. the positive results of Perron [6] as well as his negative observations in [6], pp. 269-270).

The question of the validity of Poincaré's assertion in the case $\omega = f(\infty) > 0$, say

$$(8) \qquad \omega = 1,$$

does not appear to have been treated in the literature. The chances for the truth of the assertion in this case seem to be higher than they are in the case (6). In fact, the prototypes (3) in the cases (6) and (8) are $x'' = 0$ and $x'' + x = 0$, and the former has secular solutions (i.e., it leads to a multiple elementary divisor), whereas the roots of the characteristic equation of the latter are distinct ($\lambda = \pm i$), and even the title of Perron's paper [7] emphasizes the multiple character of the roots in his case (6). It turns out, however, that Poincaré's assertion is false in the case (8) also. In fact, if

---

* In view of the superposition principle, the assertion is implied by the statement that if $f(t)$ is a real-valued continuous function and if (1) has a real-valued solution $x(t)$ for which neither $x(t)$ nor $x'(t)$ vanishes for large t, then (2) implies (5) for the case $z(t) = x(t)$ of (4).

A simple proof follows by applying l'Hôpital's rule to the function x'^2/x^2. In fact, the ratio of the derivatives of x'^2 and x^2 is $2x'x''/2xx' = x''/x$, which is $-f$, by (1), and so the assertion follows from (2).

$$f(t) = 1 + (\sin t)/(\tfrac{1}{2}t), \tag{9}$$

then (2) and (8) are satisfied (and $f(t)$, instead of being just continuous, is an entire function) but (1) fails to possess any solution $x = z(t)$ for which Riccati's quotient z'/z tends to a limit as $t \to \infty$.

In order to see this, let

$$r = r(t) = |z(t)| = (x^2 + y^2)^{\frac{1}{2}}. \tag{10}$$

Then

$$rr' = xx' + yy'; \tag{11}$$

so that, from (4),

$$w = r'/r + ic/r^2, \tag{12}$$

where c is the Wronskian constant,

$$xy' - x'y = c\ (\neq 0). \tag{13}$$

It is clear from (12) that (5) cannot hold if $r = r(t)$ is unbounded, as $t \to \infty$. This will be the case if (1) has some unbounded (real-valued) solution; that is, if (1) has some solution satisfying

$$\limsup_{t \to \infty} |x(t)| = \infty. \tag{14}$$

But Perron [8] has shown, in another connection, that (2) and (8) are compatible with (14); in fact, (9) is a simplified version of his example (cf. [10], pp. 386-387). Hence, the case (9) of (1) has no solution satisfying (5), where $\omega = 1$.

The considerations leading to these remarks will be considerably refined below.

2. The following conventions will be used: The assumption that $f(t)$ is real-valued and continuous for large t will not be repeated below. $x(t)$ and $y(t)$ denote real-valued, linearly independent solutions of (1) and $c\ (\neq 0)$ denotes their Wronskian constant (13). A solution $z = z(t)$ of (1) will be called of type (7), (13) if it is the linear combination (7) of real-valued solutions satisfying (13) with a $c \neq 0$.

(I) *Let* (1) *possess a solution* $z=z(t)$ *of type* (7), (13) *having a Riccati quotient* (4) *which satisfies* (5), *where* ω *is* (*finite and*) *real. Then*

$$(15_\omega)\qquad \operatorname*{l.u.b.}_{0<s<\infty} \Big| \int_t^{t+s} (f(u)-\omega)\,du \Big| / (1+s) \to 0 \text{ as } t\to\infty.$$

This shows the extent to which (2) can be altered when (4), (5) hold. The theorem is known in the cases that $\omega<0$ in (5) ([1], pp. 570-571) and in the case that (1) is non-oscillatory and $\omega=0$ (see [2], p. 391).

The contrast between the cases $\omega=0$ and $\omega>0$ is shown by the following theorems:

(II_0) *Let* $z=z(t)$ *be a solution of* (1) *of type* (7), (13) *for which the Riccati ratio* (4) *satisfies*

$$(16_0)\qquad w(t)\to 0 \text{ as } t\to\infty.$$

Then every solution of type (7), (13) *satisfies* (16_0).

If (16_0) is replaced by (5), where $\omega>0$, say by

$$(16_+)\qquad w(t)\to i \text{ as } t\to\infty,$$

then there is a uniqueness theorem.

(II_1) *Let* $z=z(t)$ *be a solution of* (1) *of type* (7), (13) *for which* (4) *satisfies* (16_+). *Then the solution* $z=z(t)$ *is unique up to a constant factor.*

If (16_-) means the relation which results if i is replaced by $-i$ in (16_+), then $z=z(t)$ satisfies (16_+) if and only if its complex conjugate $\bar{z}(t)$, which is also a solution of type (7), (13), satisfies (16_-).

(III_0) *Let* $z=z(t)$ *be a solution of* (1) *of type* (7), (13). *Then the conditions* (15_0) *and*

$$(17)\qquad |z(t)|\to\infty \text{ as } t\to\infty$$

are necessary and sufficient for (16_0).

The analogue for ($16_\pm$) is as follows:

(III_1) *Let* $z=z(t)$ *be a solution of* (1) *of type* (7), (13). *Then the condition* (15_1) *and the condition that*

$$(18)\qquad \lim_{t\to\infty} |z(t)| \text{ exists}$$

(*as a finite limit*) *are necessary and sufficient for* ($16_\pm$).

In this theorem, (18) can be replaced by the condition that

$$(19) \qquad \lim_{t\to\infty} |z'(t)| \text{ exists.}$$

(IV_1) *Let* $z = z(t)$ *be a solution of* (1) *of type* (7), (13). *Then* (15_1) *and* (19) *are necessary and sufficient for* ($16_\pm$).

It will be seen from the proofs that if (15_1) is assumed, then the necessary and sufficient condition (18) and/or (19) for ($16_\pm$) can be replaced by the condition that $\lim_{t\to\infty} (|z(t)|^2 + |z'(t)|^2)$ exists and is $2|c|$.

For the case $\omega = 1$ in (5), there results the following existence theorem:

(V_1) *If* (15_1) *holds, then necessary and sufficient for the existence of a solution* $z = z(t)$ *of* (1) *of type* (7), (13) *satisfying* (16_+) *is that*

$$(20) \qquad \lim_{t\to\infty} |x + ix'| \text{ exists}$$

for every (real-valued) solution $x = x(t)$ *of* (1).

It may be remarked that (20), as a sufficient condition, cannot be relaxed to the assumption that

$$(21) \qquad |x + ix'| = O(1) \text{ as } t \to \infty$$

holds for every (real-valued) solution $x = x(t)$ of (1); cf. Section 9 below. It will, however, remain undecided whether or not the sufficient condition (20) can be relaxed to the assumption that (20) holds for *some* (instead of *every*) real-valued solution $x = x(t) \not\equiv 0$.

(VI_1) *If* (1) *has a solution* $z = z(t)$ *of type* (7), (13) *satisfying* (16_+) *and if* $h = h(t)$ *is a continuous function for large positive* t *satisfying* $\int^\infty |h(t)|\, dt < \infty$, *then* $x'' + (f(t) + h(t))x = 0$ *has a solution* $z = z(t)$ *of type* (7), (13) *satisfying* (16_+).

In fact, since the solutions of (1) are bounded, as $t \to \infty$, by (III_1) and/or (IV_1), it follows from [11], p. 261, that to any solution $x = x(t)$ of (1) there corresponds a (unique) solution $x = x_0(t)$ of the last equation satisfying, as $t \to \infty$,

$$x_0(t) = x(t) + o(1), \qquad x_0'(t) = x'(t) + o(1).$$

Thus, (VI_1) is a consequence of (V_1) (for instance).

3. Proof of (I). Let $z = z(t)$ be a solution of (1) for which the Riccati ratio (4) satisfies (5). Since (I) is known when $\omega < 0$, it will be supposed that either $\omega=0$ or $\omega=1$. In view of (12), the relation (5) is equivalent to the pair of limit relations

$$r'/r\rightarrow 0 \text{ as } t\rightarrow\infty \tag{22}$$

and

$$r\rightarrow\infty \text{ as } t\rightarrow\infty \qquad \text{if } \omega=0, \tag{23_0}$$

or

$$r^2\rightarrow|c| \text{ as } t\rightarrow\infty \qquad \text{if } \omega=1. \tag{23_1}$$

The relations (23) express the fact that

$$c^2/r^4(t)\rightarrow\omega \text{ as } t\rightarrow\infty. \tag{24}$$

If (10) is differentiated twice and (1) is used in the form $x''=-fx$ and $y''=-fy$, there results the relation $-r''/r=f-(xy'-yx')^2/r^4$. In view of (13), this means that $r=r(t)$ satisfies the "linear" differential equation

$$r''+g(t)r=0, \tag{25}$$

where

$$g(t)=f(t)-c^2/r^4(t). \tag{26}$$

The equation (25) is non-oscillatory, since it possesses the solution $r=r(t)>0$. By Theorem B in [2], p. 391, a necessary and sufficient condition for (25) to possess a solution $r=r(t)$ satisfying (22) is that

$$\underset{0<s<\infty}{\text{l. u. b.}}\left|\int_t^{t+s} g(u)\,du\right|/(1+s)\rightarrow 0 \text{ as } t\rightarrow\infty. \tag{27}$$

Hence (15) follows from (24) and (26). This proves (I).

4. Proof of (II_0) and (III_0). The assertion (III_0) will be considered first. If (1) has a solution $z=z(t)$ of type (7), (13) satisfying (16_0), then (15_0) holds, by (I). In addition, the proof of (I) shows that (16_0) implies (23_0), that is, (17). Hence (15_0) and (17) are necessary for (16_0).

Conversely, if (15_0) and (17) are assumed, then (27) holds. Consequently, Theorem B of [2], p. 391, shows that the solution $r=r(t)$ of the non-oscillatory equation (25) satisfies (22). In view of (12), the relations (22) and (23_0) (that is, (17)) give (15_0). This proves (III_0).

In (II_0), it is assumed that (1) has a solution $z=z(t)$ of type (7), (13) satisfying (16_0). Thus, (15_0) holds, by (I). The assertion (III_0) implies, therefore, that a solution $z=z(t)$ of (7), (13) satisfies (16_0) if and only if (17) holds. Consequently, it is sufficient to show that if a pair of real-valued, linearly independent solutions $x = x(t)$, $y = y(t)$ satisfies (17), then any pair of real-valued, linearly independent solutions, $x = c_1x + c_2y$ and $y = c_3x + c_yy$, where $c_1c_4 - c_2c_3 \neq 0$, satisfies (17), that is, $(c_1x + c_2y)^2 + (c_3x + c_yy)^2 \to \infty$ as $t \to \infty$. But this is clear, since $c_1c_4 - c_2c_3 \neq 0$ implies that there is a const. > 0 satisfying

$$(c_1x + c_2y)^2 + (c_3x + c_4y)^2 \geqq \text{const.}\ (x^2 + y^2)$$

for arbitrary x, y. This proves (II_0).

5. **Proof of (II_1).** Suppose that (1) has a pair of solutions $z_1 = z_1(t)$, $z_2 = z_2(t)$ satisfying $z_1'/z_1 \to i$, $z_2'/z_2 \to i$ as $t \to \infty$. Then

$$z_1'/z_1 - z_2'/z_2 = (z_1'z_2 - z_1z_2')/z_1z_2 \to 0 \text{ as } t \to \infty.$$

Here the Wronskian $z_1'z_2 - z_1z_2'$ is a constant. On the other hand, $|z_1|$ and $|z_2|$ tend to finite limits, as $t \to \infty$; cf. the derivation of (23_1) from (5). Thus, the relation in the last formula line cannot hold unless $z_1'z_2 - z_1z_2' \equiv 0$, that is, unless $z_1 = \text{const.}\, z_2$. This proves ($\mathrm{II}_1$).

6. **Proof of (III_1).** The necessity of (15_1) and (18) for ($16_\pm$) follows from the assertion (I) and its proof (cf. (23_1)).

In order to prove the sufficiency of (15_1) and (18), let $z = z(t)$ be a solution of (1) of type (7), (13) satisfying (18). The proof of (III_1) will be completed first under the assumption that

(28) (15_1) and (18) imply (23_1).

It will then be shown that (28) is always true.

If (23_1) holds, then (27) follows from (15_1) and (26). In this case, (22) is a consequence of (25), (27); see [2], p. 391. Finally, ($16_\pm$) follows from (12) and (22), (23_1).

Thus, in order to complete the proof of (III_1), only the assertion (28) remains to be proved. It should be noted that if (15_1) is replaced by (2) and (8), then (23_1) is a consequence of the fact that, on the one hand, the number $N(t)$ of zeros of $y(s)$ on $0 < s < t$ satisfies

$$\pi N(t) \sim t \tag{29}$$

and, on the other hand,

$$\left|\pi N(t) - \int^{t} |c|\, ds/r^2(s)\right| \leqq \text{Const. as } t \to \infty; \tag{30}$$

cf. [4], pp. 210-211.

In order to prove (28) (without the assumptions (2) and (8)), note that, since (25) is non-oscillatory,

$$\text{either } \int^t g(s)\,ds \to -\infty \text{ or } \lim_{t\to\infty} t^{-1} \int^t \Big(\int^s g(u)\,du \Big) ds \text{ exists}$$

(as a finite limit); cf. [2], p. 390. Hence (15_1) and (26) imply that

$$1 \leqq \lim_{t\to\infty} c^2/r^4(t) \ (\leqq \infty).$$

Suppose, if possible, that (23_1) does not hold, so that

$$1 < \lim_{t\to\infty} c^2/r^4(t) \ (\leqq \infty). \tag{31}$$

Then the solution $r = r(t)$ of (25) satisfies

$$1 + (r'/r)^2 \to \lim_{t\to\infty} c^2/r^4(t) \ (\leqq \infty) \text{ as } t \to \infty. \tag{32}$$

This follows from [1], pp. 570-571, if the limit in (31) is finite and from this result and from comparison theorems (cf. [5], pp. 635-636) if the limit is infinite. In view of (31), the relations (31), (32) imply that $r'/r < 0$ for large t; in fact, $r'/r < -\text{const.} < 0$ for large t. Hence

$$r < \exp(-\text{Const.}\, t) \text{ for large } t \qquad (\text{Const.} > 0). \tag{33}$$

By (13) and Schwarz's inequality, $c^2 \leqq r^2(t)(x'^2 + y'^2)$. Consequently,

$$x'^2 + y'^2 > c^2 \exp\,(2\ \text{Const.}\, t) \text{ for large } t. \tag{34}$$

If (1) is multiplied by x and if xx'' is written as $(xx')' - x'^2$, then a quadrature leads to

$$\int^t x'^2 ds = xx' + \int^t f(s)x^2(s)\,ds + \text{const.} \tag{35_1}$$

An integration by parts gives

$$\int^t fx^2 ds = x^2F - 2\int^t xx'F ds, \tag{35_2}$$

where $F(t) = \int^t f(s)\,ds$. Since $2\,|\,ab\,| \leqq \frac{1}{2}a^2 + 2b^2$ for any pair of real numbers a, b,

$$|\,2\int^t xx'F ds\,| \leqq \tfrac{1}{2}\int^t x'^2 ds + 2\int^t F^2x^2 ds. \tag{35_3}$$

The relations (35_1), (35_2), (35_3) give

$$\tfrac{1}{2}\int^{t} x'^2 ds \leqq xx' + O(t) + O(t^3), \text{ as } t \to \infty,$$

since $F(t) = O(t)$ and $|x| \leqq r(t) = O(1)$, by (15_1) and (31), respectively. Thus

$$\tfrac{1}{2}\int^{t} (x'^2 + y'^2) ds \leqq xx' + yy' + O(t^3) \text{ as } t \to \infty.$$

Another quadrature shows that

$$\tfrac{1}{2}\int^{t} \left(\int^{u} (x'^2 + y'^2) ds\right) du \leqq \tfrac{1}{2} r^2(t) + O(t^4) \text{ as } t \to \infty.$$

In view of (31), the term $r^2(t)$ is bounded as $t \to \infty$, and so the last relation contradicts (34). This contradiction shows that the hypothesis (31) is untenable. Hence (28) follows. This completes the proof of (III_1).

Remark. In this proof of (28), the full force of (18) was not used. Correspondingly, the proof remains applicable when $\lim |z(t)|$, as $t \to \infty$, exists as a finite limit or as ∞. In fact, whether or not $r(\infty)$ $(\leqq \infty)$ exists, the proof shows that

(36) $$\liminf_{t \to \infty} c^2/r^4(t) \leqq 1 \leqq \limsup_{t \to \infty} c^2/r^4(t)\,;$$

cf. (29), (30) for the case that (2), (8) hold.

7. Proof of (IV_1). Let $\rho = \rho(t)$ denote

(37) $$\rho(t) = |z'(t)| = (x'^2 + y'^2)^{\frac{1}{2}} > 0.$$

Then a differentiation of (11) leads to $rr'' + r'^2 = \rho^2 + xx'' + yy''$. Hence, if (1) is written in the form $x'' = -fx$, $y'' = -fy$ and (25), (26) are used, it follows that

(38) $$\rho^2 = r'^2 + c^2/r^2.$$

Since $(16_\pm)$ is equivalent to (22) and (23_1), it is seen that (19) is a necessary condition for $(16_\pm)$. Thus, all that remains to be shown is that (15_1) and (19) are sufficient for $(16_\pm)$. In view of (III_1), it is sufficient to show that

(39) (15_1) and (19) imply (18).

To this end, note that

(40) $$\rho^2(\infty) = \lim_{t \to \infty} \rho^2(t) \geqq |c|.$$

In fact, it is clear from (38) and $r'^2 \geqq 0$ that $\limsup_{t\to\infty} c^2/r^2(t) \leqq \rho^2(\infty)$; so that (40) follows from the second inequality in (36).

If $\epsilon > 0$ and if t is sufficiently large, then $c^2/r^2(t) \leqq \rho^2(\infty) - 2\epsilon^2$ implies that $|r'(t)| > \epsilon$, by (38). If $t = t_0$ is sufficiently large and $c^2/r^2(t_0) \leqq \rho^2(\infty) - 2\epsilon^2$, then either $r'(t_0) > \epsilon$ or $r'(t_0) < -\epsilon$. In the former case,

$$c^2/r^2(t) \leqq \rho^2(\infty) - 2\epsilon^2,$$

hence $r'(t) > \epsilon$, for all $t \geqq t_0$. But $r'(t) > \epsilon$ for large t implies that $r(t) \to \infty$ as $t \to \infty$, and this contradicts (36). Consequently, $r'(t_0) < -\epsilon$ whenever $c^2/r^2(t_0) \leqq \rho^2(\infty) - 2\epsilon^2$ (and t_0 is sufficiently large). It follows that $c^2/r^2(t) > \rho^2(\infty) - 2\epsilon^2$ for all large t, that is, $\liminf c^2/r^2(t) \geqq \rho^2(\infty)$. Thus, by the last formula line, $\lim c^2/r^2(t)$ exists and is $\rho^2(\infty)$. Since $\rho^2(\infty) \geqq |c| > 0$, by (40), the relation (18) holds. This proves (39) and completes the proof of (IV_1).

8. Proof of (V_1). Let $z = z(t)$ be a solution of (1) of type (7), (13) satisfying (16_+). In view of (22) and (23_1),

$$rr' \to 0 \text{ and } r^2 \to c\ (> 0) \text{ as } t \to \infty. \tag{41}$$

Consider the first of these two relations and the relation (13), that is,

$$xx' + yy' = o(1) \text{ and } yx' - xy' = -c,$$

as a pair of linear relations in x' and y'. Solving for x' and y', one obtains $(x^2 + y^2)x' = -cy + o(|x|)$ and $(x^2 + y^2)y' = cx + o(|y|)$, as $t \to \infty$. Hence, in view of the second relation in (41),

$$x' = -y + o(1) \text{ and } y' = x + o(1). \tag{42}$$

Thus (20) follows from (18). Similarly,

$$\lim_{t\to\infty} |y + iy'| \text{ exists.} \tag{20 bis}$$

Finally, an arbitrary real-valued solution, $c_1x + c_2y$, of (1) satisfies

$$\begin{aligned}(c_1x + c_2y)^2 + (c_1x' + c_2y')^2 \\ = c_1^2 |x + ix'|^2 + c_2 |y + iy'|^2 + 2c_1c_2(xy + x'y').\end{aligned} \tag{43}$$

In view of (42),

$$xy + x'y' = o(1) \text{ as } t \to \infty. \tag{44}$$

Hence

$$\lim_{t\to\infty} |(c_1x + c_2y) + i(c_1x' + c_2y')| \text{ exists,} \tag{45}$$

by (43), (20) and (20 bis). This proves that the condition that (20) holds for every real-valued solution $x(t)$ of (1) is necessary for the existence of a $z=z(t)$ satisfying (16_+).

Conversely, suppose that (20) holds for every real-valued solution $x=x(t)$ of (1). In particular, (45) holds for arbitrary constants c_1, c_2 and a pair of real-valued. linearly independent solutions, $x=x(t)$ and $y=y(t)$, of (1), with $c>0$ in (13).

In view of (43), the assumption (45), for arbitrary constants c_1 and c_2, implies that $\lim (xy+x'y')$ exists. Without altering the value of c in (13), the solution $y(t)$ can be replaced by $bx+y$, where the constant b is arbitrary. Then $xy+x'y'$ becomes replaced by

$$x(bx+y)+x'(bx'+y')=b(x^2+x_1'^2)+xy+x'y'.$$

Thus, if b is suitably chosen and $bx+y$ is renamed y, then (44) holds. If the pair of relations (13) and (44), which are linear in x, y', are solved for x, one obtains $(y^2+y'^2)x=cy'+o(1)$. If the solutions x, y are replaced by const. x, y/const., then (13) and (44) remain unchanged and it can be supposed that $y^2+y'^2=c+o(1)$, $(c>0)$, after a proper choice of const., if y/const. is renamed y. Thus $x=y'+o(1)$. But then (20 bis) gives (18), hence ($16_\pm$). In fact, (16_+) holds, since $c>0$ in (13). This completes the proof of (V_1).

9. Examples to the case (8). There exist real-valued, continuous functions $f(t)$ which satisfy (2) and (8) but are such that (1) has no (complex-valued) solution $z=z(t)$ satisfying ($16_\pm$). In fact, it was pointed out above that (9) is an example of such an $f(t)$.

As mentioned above, there exist continuous $f(t)$ satisfying (2) and (8), but such that (1) has no solution $z=z(t)$ satisfying ($16_\pm$), although (21) holds for every solution $x=x(t)$ of (1). Such an $f(t)$ can be obtained by modifying the example in [5], pp. 631-632, as follows: Let $\lambda_1, \lambda_2, \cdots$ be a sequence of numbers satisfying

(46) $$\lambda_k \to 1 \text{ as } k \to \infty,$$

put $\beta_1=0$ and $\beta_k=\frac{1}{2}\pi\sum_{j=1}^{k} 1/\lambda_j$ for $k>1$, and define $f(t)$ by placing

(47) $$f(t)=\lambda_k^2 \text{ if } \beta_{k-1}\leqq t<\beta_k \qquad (k=1,2,\cdots).$$

Thus $f(t)$ is defined for $0\leqq t<\infty$. It is not continuous, but it will be clear that its discontinuities can be smoothed out (locally) without affecting the asymptotic behavior (as $t\to\infty$) of the solutions of (1).

Let $d_1, d_2, \cdots$ and $e_1, e_2, \cdots$ be the sequences of numbers

$$(48) \qquad d_k = \prod_{j=1}^{k-1} \lambda_{2j}/\lambda_{2j+1} \text{ and } e_k = \prod_{j=1}^{k-1} \lambda_{2j+1}/\lambda_{2j+2},$$

so that

$$(49) \qquad e_k = \lambda_2/(d_k\lambda_{2k}).$$

On the half-line $\beta_2 \leqq t < \infty$, define a function $x = x_1(t)$ by the formulae

$$(-1)^k d_k \cos \lambda_{2k}(t - \beta_{2k}) \text{ or } (-1)^{k+1} d_{k+1} \sin \lambda_{2k+1}(t - \beta_{2k+1})$$

according as $\beta_{2k} \leqq t < \beta_{2k+1}$ or $\beta_{2k+1} \leqq t < \beta_{2k+2}$. On the half-line, $\beta_3 \leqq t < \infty$, define $x = x_2(t)$ by the formulae

$$(-1)^k e_k \cos \lambda_{2k+1}(t - \beta_{2k+1}) \text{ or } (-1)^{k+1} e_{k+1} \sin \lambda_{2k+2}(t - \beta_{2k+2})$$

according as $\beta_{2k+1} \leqq t < \beta_{2k+2}$ or $\beta_{2k+2} \leqq t < \beta_{2k+3}$. The functions $x_1(t)$, $x_2(t)$ have continuous first derivatives for $t \geqq \beta_2$, $t \geqq \beta_3$, respectively, and, in the case (47), satisfy (1) for all $t \neq \beta_1, \beta_2, \cdots$; so that x_1, x_2 are "solutions" of (1). It is clear that

$$|x_1(\beta_{2k})| = d_k, \ |x_1'(\beta_{2k})| = 0 \text{ and } |x_2(\beta_{2k+1})| = e_k, \ |x_2'(\beta_{2k+1})| = 0.$$

Hence, (20) cannot hold for $x = x_1$ and/or $x = x_2$, respectively, unless

$$(50) \qquad \lim_{k\to\infty} d_k \text{ exists and/or } \lim_{k\to\infty} e_k \text{ exists.}$$

On the other hand, (21) holds for $x = x_1$ and $x = x_2$ (hence for every solution of (1)) if and only if

$$(51) \qquad d_k = O(1) \text{ and } e_k = O(1) \text{ as } k \to \infty.$$

It is clear that the sequence $\lambda_1, \lambda_2, \cdots$ can be chosen so as to satisfy (46) and (51) but to violate (50). Since (46)-(47) implies (2) and (8), there results an $f(t)$ for which (1) has no solution $z = z(t)$ satisfying (16_+), although (21) holds for every solution.

It is also clear that the $f(t)$ in this example can be chosen so as to satisfy $f(t) \geqq 1$ for all t (or $f(t) \leqq 1$ for all t). On the other hand, *if* $f(t)$ *satisfies* (2), (8) *and if*

$$(52) \qquad f(t) \text{ is monotone}$$

(*or just of bounded variation on* $0 \leqq t < \infty$), *then* (1) *has solutions* (7) *satisfying* $(16_\pm)$. This follows from asymptotic formulae ([11], p. 270) for the solutions of (1).

10. Examples to the case (6). As mentioned above, Perron [7] has shown that there exist functions $f(t)$ which satisfy (2) and (6) but are such that (1) has no solution $z=z(t)$ satisfying (16_0). In view of (I_0) and (II_0), an $f(t)$ has these properties if it satisfies (2), (6) and if (1) is oscillatory and possesses a solution $x=x_0(t)\not\equiv 0$ satisfying

$$x_0(t)=o(1) \text{ as } t\rightarrow\infty \tag{53}$$

or merely

$$x_0(t)=O(1) \text{ as } t\rightarrow\infty. \tag{54}$$

For $|z(t)|=\text{const.}\,|x_0(t)|$ at the zeros of a fixed solution $x=y_0(t)$, linearly independent of $x_0(t)$ (while the zeros of $y_0(t)$ tend to ∞, since (1) is assumed to be oscillatory).

It follows that the example of [3], p. 534, has the desired properties and, in addition, (52) holds (the fact that $f(t)$ is a step-function, and not continuous, is unimportant for the considerations at hand). In other words, (2) *and* (52) *do not assure the truth of Poincaré's assertion concerning* (5) *in the case* (6), although, as mentioned above, (2) and (52) are sufficient for the truth of Poincaré's assertion in the case (8).

It is curious that in the example of Perron and in the example of [3], just mentioned, (1) has a solution satisfying (54), but not (53). In fact, if (2) and (6) hold and if (1) is oscillatory (so that f is positive and non-increasing), then no solution $x=x(t)$ ($\not\equiv 0$) can satisfy (53) (since the relative maxima of $|x(t)|$ form a non-decreasing sequence); cf., e.g., [12].

It is therefore worth noting that $f(t)$ *can satisfy* (2), (6) *and have the property that* (1) *is oscillatory and has a non-trivial* ($\not\equiv 0$) *solution satisfying* (53). An example of such an $f(t)$ is given in [5], pp. 643-644. The function given there even satisfies $f(t)=O(t^{-2})$ as $t\rightarrow\infty$ (hence $\int^{\infty}|f|dt<\infty$) and is such that (53) becomes improved to $x(t)=O(t^{\frac{1}{2}-\frac{1}{2}\gamma})$ as $t\rightarrow\infty$, where γ (>1) is a fixed (but arbitrary) number.

Note that (2), (6) prevent that (53) holds for *all* solutions of (1). In fact, the same is true if (53) is relaxed to (54). This follows from (30) and from the fact that (29) must be replaced by $N(t)=o(t)$ as $t\rightarrow\infty$ in the case (6); cf. [13].

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] P. Hartman, "Unrestricted solution fields of almost-separable differential equations," *Transactions of the American Mathematical Society*, vol. 63 (1948), pp. 560-580.

[2] ———, "On non-oscillatory linear differential equations of second order," *American Journal of Mathematics*, vol. 74 (1952), pp. 389-400.

[3] ——— and A. Wintner, "On non-conservative linear oscillators of low frequency," *ibid.*, vol. 70 (1948), pp. 529-539.

[4] ——— and A. Wintner, "A criterion for the non-degeneracy of the wave equation," *ibid.*, vol. 71 (1949), pp. 206-213.

[5] ——— and A. Wintner, "Oscillatory and non-oscillatory linear differential equations," *ibid.*, vol. 71 (1949), pp. 627-649.

[6] O. Perron, "Ueber lineare Differentialgleichungen, bei denen die unabhängige Variable reell ist," *Journal für die reine und angewandte Mathematik*, vol. 142 (1913), pp. 254-270.

[7] ———, "Ueber das infinitäre Verhalten der Integrale einer linearen Differentialgleichung zweiter Ordnung, wenn die characteristische Gleichung zwei gleiche Wurzeln hat," *Sitzungsberichte der Heidelberger Akademie*, 1917, Abhandlung 9.

[8] ———, "Ueber ein vermeintliches Stabilitätskriterium," *Göttinger Nachrichten*, 1930, pp. 28-29.

[9] H. Poincaré, *Œuvres*, vol. 1, 1928.

[10] A. Wintner, "The adiabatic linear oscillator," *American Journal of Mathematics*, vol. 68 (1946), pp. 385-397.

[11] ———, "Asymptotic integrations of the adiabatic oscillator," *ibid.*, vol. 69 (1947), pp. 251-272.

[12] ———, "Stability and high frequency," *Journal of Applied Physics*, vol. 18 (1947), pp. 941-942.

[13] ———, "A criterion of oscillatory stability," *Quarterly of Applied Mathematics*, vol. 7 (1949), pp. 115-117.

ISOMORPHISMS BETWEEN PROJECTIVE UNITARY GROUPS.*[1]

By John H. Walter.

I. Introduction and Preliminaries.

1. 1. Introduction. The problem of finding the automorphisms of the classical groups and the related problem of determining isomorphisms between particular groups has been studied by Dieudonné [5] and Rickart [7, 8] among others. With regard to the projective unitary groups, this question has been unanswered except when the underlying vector spaces are over finite fields (Dieudonné [5, p. 82]; the proof given there for the proper projective unitary groups applies also to the projective unitary groups). We shall treat this problem here.

Denote by E a right vector space of dimension n over a sfield K of characteristic not 2. Assume furthermore that K possesses an involutive antiautomorphism $J: \phi \rightarrow \phi^J$. Then we can define a sesquilinear form f as a mapping $f: E \times E \rightarrow K$ such that $f(x, y)$ is additive in both x and y and $f(x\alpha, y\beta) = \alpha^J f(x, y)\beta$ where α and β are in K. The form f is said to be hermitian if $f(y, x)^J = f(x, y)$ and nondegenerate if $f(x, y) = 0$ for all y in E implies that $x = 0$. We shall assume henceforth that f is a nondegenerate herimtian sesquilinear form.[2]

The group $U_n(K, f)$ of the linear transformations of E onto itself which leave invariant the form f is called the *unitary group*. The inner automorphism $\phi \rightarrow \lambda^{-1}\phi\lambda$ of K induced by the element λ will be denoted by λ^*; so $\phi^{\lambda^*} = \lambda^{-1}\phi\lambda$. The transformation $h_\lambda: x \rightarrow x\lambda$ for all x in E is semilinear with respect to the automorphism λ^* of K and is called a *homothetic transformation*. The group of homothetic transformations is denoted by $H_n(K)$.

---

* Received June 8, 1955.

[1] This paper is based on the author's dissertation submitted for the degree of Doctor of Philosophy at the University of Michigan. The author acknowledges the guidance and encouragement given by Professor J. A. Dieudonné. In part, this work has been sponsored by the Office of Ordnance Research, U. S. Army, under Contract DA-20-018-ORD 13281.

[2] Birkhoff and von Neumann [1] have shown that if f is reflexive and not alternating, then there exists an element α in K such that αf is hermitian. The projective group associated with these forms then will be identical to the projective unitary groups.

The factor group $U_n(K,f)/H_n(K)\cap U_n(K,f)$ is the *projective unitary group* $PU_n(K,f)$.

Dieudonné [5] and Rickart [9] have shown that if certain involutions of $PU_n(K,f)$, called extremal involutions, which are defined by the way they act on E and which are associated with the 1-dimensional nonisotropic subspaces of E can be characterized strictly in terms of the multiplicative structure of $PU_n(K,f)$, then the automorphisms of $PU_n(K,f)$ can be determined and the related question of determining the isomorphisms between certain subgroups of $PU_n(K,f)$ and similar subgroups of another projective unitary group can be answered. The group-theoretical characterization of extremal involutions is our principal result.

Besides the aforesaid restriction on K, we assume, whenever we use Lemma 1. 5, that K is not the finite field $GF(3)$, a case which has been considered by Dieudonné [5]. For the sake of convenience, we furthermore assume that n is finite although the extension to the infinite dimensional case is immediate due to some remarks of Rickart [8; p. 700]. We shall ultimately restrict n to be greater than 4 but we do not assume this from the outset. Of course, we take $n \geq 2$ in all cases.

1. 2. Definitions and Notations.[3] In the study of the projective unitary group, we shall consider also the group of collineations $\Gamma L_n(K)$ which consists of all 1-1 semilinear transformations of E, and the factor group $P\Gamma L_n(K) = \Gamma L_n(K)/H_n(K)$, which is called the group of projective collineations. By the group $\Gamma U_n(K,f)$ of semisimilitudes, we understand the subgroup of $\Gamma L_n(K)$ consisting of those elements t which satisfy the equation

$$f(t(x),t(y)) = e_t f(x,y)^\tau$$

where τ is the automorphism of K associated with t and e_t is an element of K, which depends only on t and is called "the multiplier of t." It can be shown that if ϕ is in K, $\phi^{\tau J} = e_t \phi^{J\tau} e_t^{-1}$. A homothetic transformation h_λ is contained in $\Gamma U_n(K,f)$ with multiplier $\lambda^J\lambda$. The factor group

$$P\Gamma U_n(K,f) = \Gamma U_n(K,f)/H_n(K)$$

is called the group of projective semisimilitudes.

If K is commutative and J is the identity, the groups $U_n(K,f)$ and $PU_n(K,f)$ become orthogonal and projective orthogonal groups $O_n(K,f)$

---

[3] The reader is referred to Dieudonné [2] from which this paper takes its point of departure. In places where details have been omitted, the reader is referred to Dieudonné's treatise.

and $PO_n(K,f)$, respectively, which we merely consider as special cases of the unitary and projective unitary groups.

The symplectic groups will be introduced. Such a group is defined, when K is commutative, as the group of linear transformations leaving invariant an alternating bilinear form f and will be denoted by $Sp_n(K,f)$. The corresponding group of semilinear transformations, however, will still be denoted by $\Gamma U_n(K,f)$.

With the exception of the case where $U_n(K,f)$ is the orthogonal group $O_2(K,f)$, K is the finite field $GF(3)$, and f is of index 1, the center of $U_n(K,f)$ is $C_n = H_n(K) \cap U_n(K,f)$. Let Z denote the center of K; then C_n consists of those homothetic transformations h_λ, where λ is in and $\lambda^J\lambda = 1$. The set of such elements λ forms a multiplicative group of Z which we denote by C. At times, a homothetic transformation h_λ will be denoted by λ and $h_\lambda(x)$ by $\lambda(x)$.

1.3. Classification of the involutions of $PU_n(K,f)$.

We shall use the methods of linear algebra and for that reason we look upon an element t^* of $PU_n(K,f)$ as a coset in $U_n(K,f)$. In particular, if u^* is an involution[4] of $PU_n(K,f)$, then an element u of u^* will be called a *projective involution.* If u^* is an involution of $P\Gamma U_n(K,f)$ then u^* will be called a *semilinear involution* and u of $\Gamma U_n(K,f)$ in the coset u^* will be called a *projective semilinear involution.* The relation $u^2 = \gamma$ with γ in C_n characterizes projective involutions.

We shall now distinguish various types of projective involutions u, where $u^2 = \gamma$ in C_n:

(i) u is of the *first kind* if $\gamma = \lambda^2$ when λ is in Z and $\lambda^J\lambda = 1$.

(ii) u is of the *second kind* if $\gamma = \lambda^2$ when λ is in Z and $\lambda^J\lambda = -1$.

(iii) u is of the *third kind* if $\gamma = \lambda^2$ when λ is not in Z.

(iv) u is of the *fourth kind* if γ is not a square in K.

Note that if $\gamma = \lambda^2$ and λ is in Z, then $\gamma^J\gamma = 1$ implies that $\lambda^J\lambda = \pm 1$. So in this case u is either of the first or second kind. If a projective involution u is of the k-th kind, then so is every projective involution ηu, where η is in C_n. So we can say that the involution u^* of $PU_n(K,f)$ which contains u is also of the k-th kind. An extremal involution is a particular kind of involution of the first kind. Because the identity 1 can obviously be distinguished group-theoretically from other involutions, we shall exclude it from the set of involutions of the first kind and the elements of C_n from the set of projective involutions of the first kind.

---

[4] By an involution, we mean an element u^* such that $u^{*2} = 1$.

It is clear that if two involutions in $U_n(K,f)$ or two involutions in $PU_n(K,f)$ are conjugate, they are of the same kind.

1.4. Definition of the invariants ν_u^* and ω_u^*. We shall use in our characterization of extremal involutions of $PU_n(K,f)$ group-theoretical invariants similar to those introduced by Mackey [6] and effectively used by Rickart [7, 8]. Let $\mathcal{S}^*$ be a set of involutions in $PU_n(K,f)$ and define $c(\mathcal{S}^*)$ to be the set of those involutions of $PU_n(K,f)$ which commute with every element of $\mathcal{S}^*$. Denote by

$\nu(u^*, v^*)$	the number of elements in $c(c(u^*, v^*))$;
ν_u^*	the maximum of $\nu(u^*, v^*)$ for all involutions $v^* \neq u^*$ such that $u^*v^* = v^*u^*$, $u^* \neq 1^*$, and $v^* \neq 1^*$.

It is easily seen that $c(c(u^*, v^*))$ always contains 1^*, u^*, v^*, and u^*v^* when u^* and v^* are chosen as above; so $\nu_{u^*} \geqq 4$. Rickart [7, 8] showed in the cases he considered that the minimum of the invariant corresponding to ν_{u^*} was attained only for extremal involutions. In the case of the projective unitary groups, this is no longer true. So we introduce a new invariant.

Let $\mathfrak{S}^*$ denote the set of involutions of $PU_n(K,f)$ for which $\nu_{u^*} = 4$. As ν_{1^*} is not defined, 1^* is not in $\mathfrak{S}^*$. With one exception when $n = 6$, we shall show that $\mathfrak{S}^*$ contains the extremal involutions and the involutions of the third and fourth kinds. Denote by

$\omega(u^*, v^*, w^*)$	the number of elements of $c(c(u^*, v^*, w^*))$;
ω_u^*	the maximum of $\omega(u^*, v^*, w^*)$ subject to the conditions that u^*, v^*, and w^* are distinct mutually commuting involutions in $\mathfrak{S}^*$ such that $u^*v^* \neq w^*$.

The condition that $u^*v^* \neq w^*$ implies that $\omega_{u^*} \geqq 8$. We have been able to show that the invariant ω_{u^*} alone distinguishes extremal involutions when $n > 4$ and $n \neq 6$, 8, or 12. However, in order to include the cases $n = 6$, 8 and 12, we do not rely completely on this invariant in this paper.

If u^* and v^* are commuting involutions, Dieudonné [5, p. 8] has shown that for u in u^* and v in v^*, $uv = \pm vu$ and that the choice of sign is independent of the choice of representatives. It is convenient to denote the set of elements of $U_n(K,f)$ determined by a set $\mathcal{S}^*$ of $PU_n(K,f)$ by $\mathcal{S}$. Conversely, if $\mathcal{S}$ is a set of elements in $U_n(K,f)$, we write $\mathcal{S}^*$ for the set of elements in $PU_n(K,f)$ which have, as cosets, representatives in $\mathcal{S}$. Likewise, when $\mathcal{S}$ is a set of projective involutions, let $c(\mathcal{S})$ and $c(c(\mathcal{S}))$ denote the sets of projective involutions determined by $c(\mathcal{S}^*)$ and $c(c(\mathcal{S}^*))$, respectively. Then $c(\mathcal{S})$ can be characterized as the set of projective involutions which commute or anticommute with the elements of $\mathcal{S}$.

1.5. Some concepts in linear algebra. The usual concepts of orthogonality, isotropic, and nonisotropic subspaces, etc., will be used. By $S^\perp$, we mean the subspace orthogonal to a given subspace S. A 2-dimensional subspace will sometimes be called a plane.

We shall say that a 1-dimensional subspace X of a vector space E is *f-commensurable* (or in short, *commensurable*) to a 1-dimensional space Y if there exists a linear t of E mapping X onto Y such that $f(t(x), t(x)) = f(x, x)$ for all x in X. We shall say that the vector x is *f-commensurable* to the vector $t(x)$. The relation of commensurability is obviously an equivalence relation. A subspace U of E is said to be *f-commensurable* to a subspace V if $\dim U = \dim V$ and the 1-dimensional nonisotropic subspaces of U are commensurable to 1-dimensional subspaces of V.

LEMMA 1.1. *Suppose that E can be decomposed into the direct sums*

$$E = U_1 \oplus U_2 \oplus \cdots \oplus U_k \text{ and } E = V_1 \oplus V_2 \oplus \cdots \oplus V_k$$

of orthogonal subspaces such that U_i is commensurable to V_i, $i = 1, 2, \cdots, k$. Then there exists an element t in $U_n(K,f)$ such that $u(U_i) = V_i$, $i = 1, 2, \cdots, k$.

LEMMA 1.2. *Suppose that E can be decomposed into the direct sum $E = U_1 \oplus \cdots \oplus U_h \oplus V_1 \oplus \cdots \oplus V_h \oplus W$ of orthogonal subspaces such that U_i is commensurable to V_i, $i = 1, 2, \cdots, h$. Then there exists an involution u in $U_n(K,f)$ such that $u(U_i) = V_i$, $i = 1, 2, \cdots, h$ and u leaves fixed W.*

The proofs of these lemmas are obvious and will be left to the reader.

We now say that E is a *f-commensurable* (or, in short, *commensurable*) vector space if all its 1-dimensional nonisotropic subspaces are commensurable. It follows that any two nonisotropic subspaces of the same dimension which are contained in a commensurable vector space are commensurable to each other.

LEMMA 1.3. *Let $E = U \oplus V$ where U and V are orthogonal nonisotropic subspaces which are commensurable to each other. If $\dim U = \dim V \geqq 3$, E is commensurable.*

Let z be a nonisotropic vector of E not in U or V. We shall prove the lemma by showing that zK is commensurable to a nonisotropic subspace of U or V. Indeed, $z = x_1 + x_2$, where x_1 is in U and x_2 is in V. Not both x_1 and x_2 can be isotropic; for then z would be isotropic. Suppose x_1 is nonisotropic. If x_2 is nonisotropic let $Q = x_2K$ and pick y in $Q^{\perp} \cap V$ so that y is nonisotropic. If x_2 is isotropic, let Q be a nonisotropic plane contained in V and containing x_2. Then, $\dim Q^{\perp} \cap V \geqq 1$; so again pick y in $Q^{\perp} \cap V$ so that y is nonisotropic.

Let $X = x_1K$, $X' = X^{\perp} \cap U$, $Y = yK$, $Y' = Y^{\perp} \cap V$. Then

$$E = X \oplus X' \oplus Y \oplus Y'.$$

Hence by Lemma 1.2, there exists an involution u of $U_n(K,f)$ such that $u(x_1) = y\alpha$, where α is in K, and u leaves fixed X' and Y'. But then $u(z) = y_1\alpha + x_2$ is in V; so zK is commensurable to $(y_1\alpha + x_2)K$ in V.

1.6. Involutions of the first kind.[5] Let u^* be an involution of the first kind in $PU_n(K,f)$ and choose a representative u in u^*. Then $u^2 = \gamma$ and $\gamma = \lambda^2$, where λ is in C. The projective involution $\lambda^{-1}u$ is easily seen to be an involution in u^*; so we may consider u to be an involution. Then E is the direct sum $U^+ \oplus U^-$ of the eigenspaces of u, where $u(x) = \pm x$ according as x is in U^+ or U^-, respectively. The subspaces U^+ and U^- are orthogonal and nonisotropic (Dieudonné [4; p. 19]). If $\dim U^+ = p$, u is said to be a *p-involution.* As both u and $-u$ are in u^*, we can always choose u so that $p \leqq [n/2]$, where $[r]$ designates the greatest integer less than or equal to a rational number r. With this choice of representative, we also call u^* a *p-involution.* Those linear involutions for which $p = 1$ or $p = n - 1$ are called *minimal involutions* and the corresponding cosets are called *extremal involutions.*

An element t in $\Gamma L_n(K)$ commutes with an involution u if and only if t leaves invariant the eigenspaces of u. If u and v are commuting involutions of $U_n(K,f)$ with eigenspaces $U^{\pm}$ and $V^{\pm}$, respectively, then $E = P_1 \oplus P_2 \oplus P_3 \oplus P_4$, where $P_1 = U^+ \cap V^+$, $P_2 = U^+ \cap V^-$, $P_3 = U^- \cap V^+$, and $P_4 = U^- \cap V^-$. The subspaces P_i are orthogonal and nonisotropic when they are not zero. If t in $\Gamma L_n(K)$ anticommutes with u, then t maps the eigenspaces U^+ and U^- onto one another. In this case, $\dim U^+ = \dim U^- = n/2 = m$. That is, *no element t of $\Gamma L_n(K)$ can anticommute with an involution u of $U_n, K, f)$ except if u is a m-involution.*

---

[5] Dieudonné [2] has characterized the centralizers of projective involutions in $\Gamma U_n(K,f)$. The next four sections are an adaptation of this characterization to our case.

If two 1-involutions commute, their 1-dimensional eigenspaces are orthogonal. Thus a maximal set of commuting 1-involutions consists of n involutions whose 1-dimensional eigenspaces are mutually orthogonal. When $n > 2$, no element of $U_n(K,f)$ anticommutes with a 1-involution. Consequently a maximal set of commuting 1-involutions of $U_n(K,f)$ determines a maximal set of commuting extremal involutions of $PU_n(K,f)$ which always contains n involutions. The product of p mutually commuting 1-involutions of $U_n(K,f)$ is easily seen to be a p-involution. Therefore, the product of p mutually extremal involutions is a p-involution or $(n-p)$-involution according as $p \leqq [n/2]$ or $n-p \leqq [n/2]$.

When E is commensurable, it is a direct consequence of Lemma 1.1 that, for each p, all p-involutions in $U_n(K,f)$ are conjugate. Then, for each p, all p-involutions of $PU_n(K,f)$ are conjugate.

1.7. Involutions of the second kind. Let u^* be an involution of the second kind in $PU_n(K,f)$ and choose a representative u in u^*. Then $u^2=\gamma$, where $\gamma=\lambda^2$, λ is in Z, and $\lambda^J\lambda=-1$. Then $u'=\lambda^{-1}u$ is a linear involution of $\Gamma U_n(K,f)$ with multiplier -1 and $E=U^+ \oplus U^-$, where U^+ and U^- are the eigenspaces of u' and of U for any u in u^*. The subspaces U^+ and U^- are totally isotropic of dimension $m=n/2$; consequently n must be even [3, p. 225]. The conditions for t in $\Gamma L_n(K)$ to commute with u are the same as for involutions of the first kind. If v is a projective involution of the first or second kind, then we have $E=P_1 \oplus P_2 \oplus P_3 \oplus P_4$, where the subspaces P_i are defined in Section 1.6.

1.8. Involutions of the third kind. We shall consider the following generalization of projective involutions of the third kind, which arises in the subsequent developments. Let u be a projective semilinear involution of $\Gamma U_n(K,f)$ relative to an involutive automorphism σ with multiplier e such that $u^2=\gamma$ in C_n, where $\gamma=\lambda\lambda^\sigma$ for some λ in K. In case $\lambda^\sigma=\lambda$, we furthermore assume that λ is not in Z. Then $u'=\lambda^{-1}u$ is a semilinear involution of $\Gamma U_n(K,f)$ relative to the automorphism $\sigma'=(\lambda^{-1})^*\sigma=\sigma\lambda^*$ and with multiplier $e'=(\lambda\lambda^J)^{-1}e$. It follows that $\sigma'^2=1$ and $\sigma'\neq 1$. Let K_1 be the fixed subfield of K by σ'. Then there exists an element θ in K such that $\theta^{\sigma'}=-\theta$ and K is the direct sum of the two K_1-spaces: $K=K_1 \oplus K_1\theta$. Thus E can be considered as a right vector space E^* over K_1 and $\dim_{K_1}E^*=2n$. As a transformation of E^*, u' is a linear involution. So $E^*=U^+ \oplus U^-$, where U^+ and U^- are the eigenspaces of u'. Furthermore $U^-\theta=U^+$; so $\dim_{K_1}U^+=\dim_{K_1}U^-=n$ and $E^*=U^+ \oplus U^+\theta$. Thus a K_1-basis for U^+ is also a K-basis for E. We shall sometimes denote U^+ as E_1.

Now we associate with f another nondegenerate sesquilinear form defined on E_1. In general, we have $\phi^{J\sigma'} = e'^{-1}\phi^J e'$ for ϕ in K_1. Dieudonné [2; chap. I, § 14] shows that e can always be taken to be 1 by multiplying the form f by an element d^{-1} of K such that $e' = d^{1-\sigma}$ and $d^J = \pm d$. In fact, if $e' \neq -1$, it suffices to take $d = 1 + e'$. If $e' = -1$ and there exists an element ψ of $K_1\theta$ such that $\psi^J \neq -\psi$, take $d = \psi + \psi^J$. In both these cases $d^J = d$ and so the form $g = d^{-1}f$ is a hermitian sesquilinear form relative to the antiautomorphism Jd^*.

However, if $\psi^J = -\psi$ for all elements ψ of $K_1\theta$ and if ϕ is any element of K_1, we have $(\phi\psi)^J = -\phi\psi = \psi^J\phi^J = -\psi\phi^J$. So the restriction of J to K_1 is the automorphism $\phi \rightarrow \psi^{-1}\phi\psi$ where ψ is any element of $K_1\theta$; hence K_1 is commutative.[6] Now take any element of $K_1\theta$ for d. Then $d^{1-\sigma'} = -1 = e$. But now g is a skew hermitian sesquilinear form relative to the antiautomorphism $T = Jd^*$. Furthermore, since d^2 is in K_1, $\phi^T = d^{-1}\phi^J d = d^{-2}\phi d^2 = \phi$ for ϕ in K_1 and thus the restriction of T to K_1 is the identity.

Now we have that $\Gamma U_n(K,g) = \Gamma U_n(K,f)$ and $U_n(K,g) = U_n(K,f)$. In $\Gamma U_n(K,g)$, u' has multiplier 1; so if x and y are in E_1,

$$g(x,y) = g(u'(x), u'(y)) = g(x,y)^{\sigma'}$$

and $g(x,y)$ is in K_1. Thus we can restrict g to a sesquilinear form f_1 such that $f_1 : E_1 \times E_1 \rightarrow K_1$. In case g is hermitian, f_1 is hermitian and we form the groups $U_n(K_1,f_1)$ and $\Gamma U_n(K_1,f_1)$ acting on E_1.

In case g is skew-hermitian, f_1 is alternating and we form the group $Sp_n(K_1,f_1)$. Let x and y be in E_1 such that $g(x,y) = f_1(x,y) = 1$. Then $g(x,y) = g(u(x),u(y)) = \lambda^T\lambda g(x,y)^{\lambda^*}$. So $\lambda^T\lambda = 1$. In case $\lambda^\sigma = \lambda$, λ is in K_1 and we have $\lambda^2 = 1$. But we have excluded the case that $\lambda^\sigma = \lambda$ and λ is in Z. Hence f_1 can be alternating only if $\lambda^\sigma \neq \lambda$; and thus when u is linear, this case is excluded.

An element t of $\Gamma U_n(K,f)$ relative to an automorphism τ, where $\lambda^\tau = \lambda$ and with multiplier e_t which commutes with u commutes also with u'. It then follows that $\sigma'\tau = \tau\sigma'$; so $K_1{}^\tau = K_1$. Thus t can be restricted to a semilinear transformation t_1 acting on E_1. Furthermore, t is in $\Gamma U_n(K,g)$ with multiplier $e_t' = d^{-1}e_t d$. But then

$$g(u't(x), u't(y)) = e_t'^{\sigma'} g(x,y)^{\sigma'\tau} = g(tu'(x), tu'(y)) = e_t g(x,y)^{\sigma'\tau}.$$

Since $\tau\sigma' = \sigma'\tau$, $e_t'^{\sigma'} = e_t'$ and e_t' is in K_1. Thus the restriction t_1 of t to E_1 is in $\Gamma U_n(K_1,f_1)$ or $SP_n(K_1,f_1)$ relative to the restriction of τ to K_1, as the case may be. We shall say that t_1 *represents* t in $\Gamma U_n(K_1,f_1)$.

---

[6] K can be shown to be either commutative or a sfield of generalized quaternions in this case.

In general, an element of $\Gamma U_n(K_1, f_1)$ cannot be extended to an element t of $\Gamma U_n(K, f)$. But if t_1 is in $U_n(K_1, f_1)$ or $Sp_n(K_1, f_1)$ this can be done by setting $t(y) = t_1(x)\theta$ for y in U^-, where x is the vector in $U^+ = E_1$ such that $y = x\theta$.

To summarize we have the following remark.

An element t of $\Gamma U_n(K, f)$ which is relative to an automorphism τ such that $\lambda^\tau = \lambda$ and which commutes with the projective semilinear involution u is represented by its restriction t_1 to E_1 in $\Gamma U_n(K_1, f_1)$. An element of $U_n(K_1, f_1)$ or $Sp_n(K_1, f_1)$ represents an element of $U_n(K, f)$ which commutes with u.

Let Z_1 be the center of K_1. For ϕ in K_1, we have $\phi^{\theta^*\sigma'} = \theta^{-\sigma'}\phi^{\sigma'}\theta^{\sigma'} = \phi^{\theta^*}$; so θ^* leaves K_1 and Z_1 invariant. Since θ^2 is in K_1, the restriction of θ^* to Z_1 is an involutive automorphism of Z_1. Thus either every element of Z_1 commutes with θ (in case the restriction of θ^* is the identity) or by Galois theory $Z_1 = Z_0 \oplus Z_0\lambda'$, where $\lambda'\theta = -\theta\lambda'$ and every element of Z_0 commutes with θ.

Let Z' be the subfield of Z_1 which is left fixed by θ^*. Since $K = K_1 \oplus K_1\theta$, $Z' \subseteq Z$. On the other hand, $K_1 \cap Z \subseteq Z'$. So as $Z' \subseteq K_1$, $Z' = K_1 \cap Z$. Therefore if θ^* is the identity on Z_1, $Z_1 = K_1 \cap Z$. If θ^* is order 2 on Z_1 with fixed subfield Z_0, $Z_0 = K_1 \cap Z$.

In the important case where u is a projective involution of the third kind in $U_n(K, f)$, $\sigma' = (\lambda^{-1})^*$ and so λ is in Z_1. But $\lambda^{\theta^*} = -\lambda$. So θ^* is not the identity on Z_1. We can take $\lambda' = \lambda$. Then $Z_1 = Z \oplus Z\lambda$ as $K_1 \cap Z = Z$; that is, $Z_1 = Z(\lambda)$.

In general, when σ' leaves Z_1 fixed and θ^* does not, we have $Z_1 = Z(\lambda')$. In this case, K_1 is an interior Galois subsfield of K, and, according to a result by Dieudonné [3, p. 182], $K = K_1 \oplus K_1\psi$, where ψ is an element of K whose square $\psi^2 = \alpha$ is in Z. However, this implies that $\psi^{\sigma'} = -\psi$. Indeed $\psi^{\sigma'} = a + b\psi$ with a and b in K_1. Therefore $\psi = \psi^{\sigma'2} = a(b+1) + b^2\psi$. Thus $b = \pm 1$. If $b = 1$, then $a = 0$ and $\psi^{\sigma'} = \psi$, which contradicts the fact that ψ is not in K_1. Therefore, $\psi^{\sigma'} = a - \psi$ and

$$(\psi^{\sigma'})^2 = a^2 + \psi^2 - (\psi a + a\psi) = (\psi^2)^{\sigma'} = \alpha.$$

This implies $a = 0$ and $\psi^{\sigma'} = -\psi$. Since θ was chosen arbitrarily to satisfy the condition $\theta^{\sigma'} = -\theta$, we can henceforth assume that $\theta = \psi$ so that $\theta^2 = \alpha$ is in Z and $(\theta^*)^2 = 1$.

Now we consider the case that t in $\Gamma U_n(K, f)$ anticommutes with u, where u is now a projective involution of the third kind in $U_n(K, f)$. Let τ

be the automorphism of K associated with t and let e_t be the multiplier of t. Assume $\lambda^\tau = \lambda$. Then t commutes with $h_{\lambda^{-1}}$ and hence anticommutes with u'. As $\theta^{\sigma'} = -\theta$, $h_\theta u' = -u' h_\theta$. So $t' = h_\theta t$ is a projective semilinear involution of $\Gamma U_n(K,f)$ relative to the automorphism $\tau\theta^*$ and $t'u' = u't'$. The multiplier of t' in $\Gamma U_n(K,g)$ is $d^{-1}\theta^J e_t \theta d^{\tau\theta^*}$ which again must be in K_1 by the argument given previously in this section. Furthermore $K_1^{\tau\theta^*} = K_1$; so t' can be restricted to a transformation t_1 of $\Gamma U_n(K_1, f_1)$ acting on E_1. We say that in this case t_1 *represents* t.

In summary: *Let t be a projective semilinear involution of $\Gamma U_n(K,f)$ relative to an automorphism τ of K. Then if $ut = -tu$ and $\lambda^\tau = \lambda$, t is represented by a transformation t_1 in $\Gamma U_n(K_1, f_1)$, which is semilinear with respect to the automorphism $\tau\theta^*$.*

1. 9. Involutions of the fourth kind. Let u^* be an involution of the fourth kind in $PU_n(K,f)$ and choose a representative u in u^*. Then $u^2 = \gamma$ in C_n, where now γ is not a square in K. Again we consider a generalization. Let u be a projective semilinear involution in $\Gamma U_n(K,f)$ relative to an involutive automorphism σ of K and with multiplier e such that $u^2 = \gamma$, where γ is in C and is not of the form $\lambda\lambda^\sigma$ for λ in K.

Dieudonné [2, chap. I, § 13] has shown that there exists a quadratic extension L of K obtained by adjoining a square root ρ of γ such that $\phi\rho = \rho\phi^\sigma$ for ϕ in K. (Obviously L is not commutative when $\sigma \neq 1$, even though K may be. So the consideration of sfields is necessary at this point.) Now E can be considered as a right vector space F over L by setting $x\rho = u(x)$ for x in E. The L-dimension of F is $m = n/2$. Consequently n must be even, if this type of projective involution is to exist in $\Gamma U_n(K,f)$. The antiautomorphism J can be extended to L by setting $\rho^J = \rho e^\sigma \gamma^{-1} = e\rho^{-1}$. Corresponding to f, we define a new nondegenerate hermitian sesquilinear form g by setting $g(x,y) = f(x,y) + \rho f(x, u(y))^\sigma \gamma^{-1}$. In $\Gamma U_n(L,g)$ u is the homothetic transformation h_ρ with multiplier $e = \rho^J \rho$.

It is easily seen that if t is an element of $\Gamma U_n(K,f)$ which is relative to an automorphism τ of K and which commutes with u, then t can be considered to be in $\Gamma U_m(L,g)$ relative to the extension of the automorphism τ to L obtained by setting $\rho^\tau = \rho$. Conversely, if t is an element of $\Gamma U_n(L,g)$ which is relative to an automorphism τ such that $K^\tau = K$, then t can be considered as an element of $\Gamma U_n(K,f)$ which is relative to the restriction of τ to K and which commutes with u.

Let σ_0 be the automorphism of L leaving K fixed which is defined by the relation $\sigma_0 : \rho \rightarrow -\rho$. Clearly σ_0 commutes with any automorphism τ

of L such that $K^\tau = K$ and $\rho^\tau = \eta\rho$ with η in K. Therefore, if t is an element of $\Gamma U_n(K, f)$ which is relative to an automorphism τ of K and which anticommutes with u, t can be considered to be in $\Gamma U_m(L, g)$ relative to the automorphism $\sigma_0\tau = \tau\sigma_0$, where now τ is extended to L by setting $\rho^\tau = \rho$. Conversely, any element t of $\Gamma U_m(L, g)$ which is relative to an automorphism of the form $\sigma_0\tau$, where $\rho^\tau = \rho$ and $K^\tau = K$, can be considered as an element of $\Gamma U_n(K, f)$ which is relative to the restriction of τ to K and which anticommutes with u.

Denote the center of L by Z_L. If $\phi = \xi + \rho\eta$ is in Z_L with ξ and η in K, the condition that ϕ should commute with every element of K implies that ξ and η are in Z. Hence $Z \subseteq Z_L \subseteq Z(\rho)$. Now if $\sigma = 1$, ρ is in $Z_L = Z(\rho)$. However, if $\sigma \neq 1$, $Z_L = Z$ as $Z(\rho)$ is a quadratic extension of Z.

In Section 4.2, we make reference to the case where $U_n(K, f)$ is the orthogonal group $O_n(K, f)$ and u is a projective involution of the fourth kind in $O_n(K, f)$. Then K is commutative and J is the identity by definition of $O_n(K, f)$. If $\gamma = u^2$, we have now $\gamma^J\gamma = \gamma^2 = 1$. As γ is not a square in K, $\gamma = -1$, $\rho = i$ and $L = K(i)$, where $i = \sqrt{-1}$. In this case, the extension of the identity automorphism J to L is obtained by setting $i^J = i^{-1} = -i$.

1. 10. A result on fields. The following remark is a special case of a well-known result on fields.

Let Z' be the splitting field of an irreducible equation $x^2 - \alpha = 0$ over a field Z. Then if β is an element of Z' such that β^2 is in Z, either β is in Z or $\beta = \lambda\rho$, where ρ is a root of $x^2 - \alpha = 0$ and λ is in Z.

1. 11. Some preliminary lemmas. The lemmas of this section apply to sfields K which will be identified with the sfield, subsfields, or extension sfields which we associate with the projective unitary group under investigation in the following sections.

LEMMA 1.4. *If K is not of characteristic 2, then for every pair of orthogonal nonisotropic vectors x and y and for any α in K, at least one of the vectors $x + y(\alpha - 1)$, $x + y\alpha$, and $x + y(\alpha + 1)$ is nonisotropic.*

Clearly, as x and y are nonisotropic, f is not alternating and hence is hermitian. Compute first

$$f(x + y\beta, x + y\beta) = f(x, x) + \beta^J f(y, y)\beta. \tag{1.1}$$

The lemma is trivial if $\alpha = 0$ or $\alpha = \pm 1$. So suppose $\alpha \neq 0$ and $\alpha \neq \pm 1$

and that $x + y\alpha$ and $x + y(\alpha \pm 1)$ are all isotropic. Then (1.1) vanishes for $\beta = \alpha$ and $\beta = \alpha \pm 1$. Setting $\xi = f(y,y) \neq 0$, we have

$$(1.2) \qquad -f(x,x) = \alpha^J \xi \alpha = (\alpha \pm 1)^J \xi (\alpha \pm 1).$$

From this it follows that $\xi = \pm (\alpha^J \xi + \xi\alpha)$, which is clearly impossible as $\xi \neq 0$. This proves the lemma.

Let r be a projective semilinear involution in a group of semisimilitudes $U_n(K,f)$ relative to an involutive automorphism σ of K such that $r^2 = \epsilon$ and ϵ is in Z. Let K_0 be the subsfield of K left fixed by σ. We must assume that $K_0 \neq GF(3)$ in the following lemma. In the applications of this lemma, this is assured, of course, if $\sigma = 1$ or if K is not commutative, whence K and K_0 are infinite. In the remaining cases, the sfield K_0 will be identified with the sfield of the projective unitary group which we investigate; so the assumption that this latter sfield is not the field $GF(3)$ is all that will be required in order to apply this lemma.

LEMMA 1.5. *Let S be a nonisotropis subspace of dimension $s \geqq 2$. Assume that the characteristic of K is not 2 and that $K_0 \neq GF(3)$. If, for every nonisotropic vector x in S, $r(x) = x\delta_x$ where δ_x is in K, then $r(x) = x\delta$ for every x in S where δ is independent of x. Furthermore $\sigma = \delta^*$.*

It is an immediate consequence of this lemma that if $\sigma = 1$, δ is in Z. In all cases σ leaves Z fixed.

Proof. Let x_0 be an arbitrary, nonisotropic vector in S. Then $r(x_0) = x_0\delta_{x_0}$. Let $\delta = \delta_{x_0}$. Let K_1 be the subsfield of K left fixed by $\pi = \delta^*\sigma$. Then $r^2 = \epsilon$ implies that $\delta\delta^\sigma = \epsilon$, which is in Z. From this it follows that $\pi^2 = 1$. Then $[K:K_1] = 1$ or 2 according as $\pi = 1$ or not. This means that $K_1 \neq GF(3)$. For, if K is infinite, this is clearly the case; if K is finite, it is commutative and $\pi = \sigma$ whence $K_1 = K_0 \neq GF(3)$.

Let x_0' be a nonisotropic vector of S orthogonal to x_0. Then $r(x_0') = x_0'\delta_{x_0'}$. Set $\delta = \delta_{x_0'}$. By Lemma 1.4, there exists an element α in K_1 such that $x_0 + x_0'\alpha$ is also nonisotropic. Because $K_1 \neq GF(3)$, we can choose $\alpha \neq 0$. Then $r(x_0 + x_0'\alpha) = (x_6 + x_0'\alpha)\delta_{x_0+x_0'\alpha}$. On the other hand $r(x_0 + x_0'\alpha) = r(x_0) + r(x_0')\alpha^\sigma$. So $x_0(\delta_{x_0+x_0'\alpha} - \delta) = x_0'(\delta'\alpha^\sigma - \alpha\delta_{x_0+x_0'\alpha})$. Because x_0 and x_0' are linearly independent relative to K, it follows that $\delta_{x_0+x_0'\alpha} = \delta$ and then that $\delta'\alpha^\sigma = \alpha\delta$. But $\alpha^\pi = \alpha$ as α is in K_1. So $\alpha^\sigma = \alpha^{\delta^*} = \delta^{-1}\alpha\delta$ and $\alpha\delta = \delta\alpha^\sigma$. Thus $\delta' = \delta$.

As x_0' was chosen to be an arbitrary nonisotropic vector orthogonal to x_0, $\delta_{x_0'\phi} = \delta$ for ϕ in K. So $r(x_0'\phi) = x_0\delta\phi^\sigma = x_0\phi\delta$. Thus $\phi^\sigma = \delta^{-1}\phi\delta$ and

$\sigma=\delta^*$. Let $x_0, x_1, \cdots, x_{s-1}$ form an orthogonal nonisotropic basis for S. By the above argument $r(x_i)=x_i\delta$. Then for

$$x=\sum_{i=0}^{s-1} x_i\alpha_i \text{ in } S,\ r(x)=\sum_{i=0}^{s-1} r(x_i)\alpha_i^{\sigma}=x\delta,$$

which completes the proof of the lemma.

LEMMA 1.6. *Let S be a subspace of E of dimension greater than 1 and let t be an element of $\Gamma U_n(K,f_1)$ relative to an automorphism τ of K. Then if $t(x)=x\delta_x$ with δ_x in K for every vector x of S, $t(x)=x\delta$, where δ is in K and is independent of x in S and τ is the inner automorphism δ^* of K.*

The proof is an obvious simplification of the proof of Lemma 1.5. This lemma includes the case that f_1 is alternating.

II. Involutions of the First Kind.

Henceforth we shall assume that $K \neq GF(3)$.

In this part u^* will represent an involution of the first kind in $PU_n(K,f)$. The usual restrictions on K shall hold. In accordance with Section 1.6, we choose the representative u of u^* which is a p-involution with $p \leqq [n/2]$. Then $E=U^+ \oplus U^-$, where U^+ and U^- are the eigenspaces of u.

LEMMA 2.1. *Assume $n>2$. If u^* is an extremal involution and v^* an involution commuting with u^*, then a projective involution v in v^* must commute with any projective involution u in u^* and v^* is either of the first or third kind.*

That v in v^* must commute with u in u^* is the result at the end of Section 1.6. Thus $v(U^+)=U^+$. If x is in U^+, $v(x)=x\delta_x$. If $v^2=\beta$, then $\beta=\delta_x^2$. If δ_x is in Z, then $\delta_x^J\delta_x f(x,x)=f(v(x),v(x))=f(x,x)\neq 0$. So $\delta_x^J\delta_x=1$. This shows that v and v^* are either of the first or third kind.

2.1. Extremal involutions commuting with involutions of the first kind. Assume $n>3$. Let u^* be an extremal involution and let $v^*\neq u^*$ be an involution of the first kind commuting with u^*. Let v be the p-involution in v^* with $p\leqq[n/2]$ and with eigenspaces V^+ and V^-. We then have $E=P_1\oplus P_2\oplus P_3\oplus P_4$, where the subspaces P_i are orthogonal and nonisotropic or 0-dimensional. As $\dim U^+=1$ one of $P_1=U^+\cap V^+$ or $P_2=U^+\cap V^-$ contains only the 0 vector. As $u^*\neq v^*$, $u\neq v$, and $U^-\neq V^-$. So $U^-=P_3\oplus P_4$, where $\dim P_3>0$ and $\dim P_4>0$. Thus E is decomposed into the direct sum of three orthogonal nonisotropic subspaces U^+, P_3 and P_4.

As $\dim E > 3$, at least one of the P_i is of dimension greater than 1. Let x be a nonisotropic vector in P_j where $\dim P_j > 1$. Let w be the 1-involution with subspaces $W^+ = xK$ and $W^- = (xK)^\perp$; w is in $U_n(K,f)$ by Section 1.6. As W^+ is in P_j, it is left invariant by u and v. Then $W^- = (W^+)^\perp$ is left invariant by u and v, and, consequently, u and v commute with w. That is, w is in $c(u,v)$; so any r in $c(c(u,v))$ must commute with w, as w is a 1-involution. Then $r(x) = x\delta_x$ and this must hold for every nonisotropic vector x in P_j. Consequently by Lemma 1.5, $r(x) = x\delta_j$ with δ_j in Z.

Furthermore, r must commute with u and commute or anticommute with v, as u and v are in $c(u,v)$. However, r cannot anticommute with v, for then would $r(P_1) = P_2$, whereas $\dim P_1 \neq \dim P_2$. So r must leave all P_i invariant. Consequently if $\dim P_k = 1$ and x is in P_k, $r(x) = x\delta_x$. But $r^2(x) = x\epsilon$ for all x in E with ϵ in C. Therefore $\delta_x^2 = \delta_j^2$. As ϵ can have but two square roots in Z and hence in K, $\delta_x = \pm\delta_j$ and δ_x is in Z. Set $\delta_k = \delta_x$. Then as δ_k is in Z, $r(x) = x\delta_k$ for all x in P_k where $\delta_k = \pm\delta_j$. If $\dim P_k > 1$, the above argument for P_j shows $r(x) = x\delta_k$ for x in P_k and again $\delta_k = \pm\delta_j$. There are thus two choices in sign for $r(x)$ when x is in the two nonzero subspaces $P_k \neq P_j$. There are thus but four possibilities for r, each of which is a representative of 1^*, u^*, v^*, or u^*v^*. Therefore $\nu(u^*, v^*) = 4$.

2.2. Extremal involutions commuting with involutions of the third kind. Assume $n > 2$. Let u^* be an extremal involution. Let v^* be an involution of the third kind and let v be in v^*. Then $v^2 = \beta$ is in C_n, where $\beta = \mu^2$ and μ is not in Z. With regard to v, form the group $U_n(K_1, f_1)$ acting on the space E_1 where K_1 is the centralizer of μ in K (cf. Section 1.8). The center of Z_1 of K_1 is $Z(\mu)$. As v_1 is linear, the form f_1 is not alternating (cf. Section 1.8). Then u is represented in $U_n(K_1, f_1)$ by its restriction u_1 to E_1; and in $U_n(K_1, f_1)$, u_1 is a 1-involution with eigenspaces $U_1^+ = U^+ \cap E_1$ and $U_1^- = U^- \cap E_1$. For the remainder of this section, we assume merely that u_1 is of the first kind in $U_n(K_1, f_1)$. This generalization is needed in Part III.

The K_1-subspaces U_1^+ and U_1^- are orthogonal and nonisotropic (with respect to f_1), at least one of which, say U_1^+, is of K_1-dimension greater than 1 as $n > 2$. Let x be a nonisotropic vector in U_1^+ and w_1 be the 1-involution of $U_n(K_1, f_1)$ with subspaces $W_1^+ = xK_1$ and $W_1^- = (xK_1)^\perp$. Then w commutes with u_1 and its extension w to E is a 1-involution of $U_n(K,f)$ commuting with u and v. Consequently w is in $c(u,v)$ and r in $c(c(u,v))$ must commute with it.

Let θ be the element of K chosen in Section 1.8. so that $\theta^{\mu^*} = -\theta$. Then r or $r' = h_\theta r$ commutes with $v' = \mu^{-1}v$. Since $h_\theta w = wh_\theta$, r' as well as r commutes with w. Then the restriction r_1 of r or r' to E_1 commutes with w_1. Therefore $r_1(W_1{}^+) = W_1{}^+$ and $r_1(x) = x\delta_x$. If r commutes with v', then clearly r_1 is a projective involution of $U_n(K_1, f_1)$ such that $r_1{}^2 = \epsilon$ where $\epsilon = r^2$. If r anticommutes with v', then r_1 is a projective semilinear involution relative to the automorphism θ^* of K_1 and $r_1{}^2 = \epsilon\theta^2$. But θ was chosen in Section 1.8 so that $\theta^2 = \alpha$ is in Z. So $r_1{}^2 = \epsilon\alpha$ is in Z and $(\theta^*)^2 = 1$. In both cases, r_1 satisfies the hypothesis of Lemma 1.5. Consequently $r_1(x) = x\delta$ with δ in K_1 for all x in $U_1{}^+$. Furthermore, since θ^* is not an inner automorphism of K_1, r_1 cannot be semilinear with respect to θ^*. Consequently r_1 is a projective involution in $U_n(K_1, f_1)$ and δ is in $Z(\mu)$. As $r_1{}^2 = \epsilon$, $\delta^2 = \epsilon$, where ϵ is in Z. But then by Section 1.10, we have that δ is in Z or $\delta = \delta_0\mu$ with δ_0 in Z.

If $\dim U_1{}^- > 1$, the same argument shows that $r(x) = x\delta^-$ with δ^- in $Z(\mu)$ for x in $U_1{}^-$. Again $(\delta^-)^2 = \delta^2$; so $\delta^- = \pm\delta$. If $\dim U_1{}^- = 1$, then as in Section 2.1, $r_1(U_1{}^-) = U_1{}^-$ and $r(x) = x\delta_x$ for x in $U_1{}^-$. But $\delta_x{}^2 = \delta^2$; so $\delta_x = \pm\delta$ and is in $Z(\mu)$. This implies that δ_x is independent of x in $U_1{}^-$; so we set $\delta_x = \delta^-$.

As r_1 is in $U_n(K_1, f_1)$ and $U_1{}^+$ and $U_1{}^-$ are nonisotropic, $\delta^T\delta = 1$. If δ is in Z, then $\delta^T = \delta^J$. The extension of r_1 to E is r and is seen to be in 1^* or u^* according as $\delta^- = \delta$ or $\delta^- = -\delta$. If $\delta = \delta_0\mu$, by looking at $v_1{}^{-1}r_1$, we similarly conclude that $\delta_0{}^T = \delta_0{}^J$ and that the extension of $v_1{}^{-1}r_1$ to E is $v^{-1}r$ and is in 1^* or u^*. So r is in v^* or u^*v^*. This yields

THEOREM 2.1. *If the dimension of E is greater than 3, and u^* is an extremal involution in $PU_n(K, f)$, $\nu_{u^*} = 4$.*

2.3. Nonextremal involutions of the first kind.

THEOREM 2.2. *If the dimension n of E is greater than 4 and not 6, then $\nu_{u^*} > 4$ for nonextremal involutions of the first kind in $PU_n(K, f)$. If $n = 6$, $\nu_{u^*} > 4$ for all 2-involutions u^*; and if, furthermore, E is not commensurable, $\nu_{u^*} > 4$ for all 3-involutions.*

First take $n > 4$ and $n \neq 6$. Let u^* be a p-involution of $PU_n(K, f)$ with $1 < p \leqq [n/2]$ and let u be the p-involution of $U_n(K, f)$ contained in u^*. We shall exhibit an involution $v^* \neq u^*$ commuting with u^* such that $\nu(u^*, v^*) > 4$. Let P_1 be a 1-dimensional nonisotropic subspace of the eigenspace U^+ of u and P_3 a 2-dimensional nonisotropic subspace of the eigenspace

U^- of u. Set $V^+ = P_1 \oplus P_3$, $V^- = (V^+)^\perp$, $P_2 = U^+ \cap V^-$, and $P_4 = U^- \cap V^-$. Then $\dim P_2 = p - 1 > 0$ and $\dim P_4 = n - p - 2$. Since $p \leqq [n/2]$ and $n \geqq 5$, $\dim P_4 > 0$. So P_1, P_2, P_3, and P_4 are orthogonal nonisotropic subspaces whose direct sum is E. Let v be the 3-involution with positive and negative eigenspaces V^+ and V^-, respectively; then clearly v commutes with u.

Let r_i, $i = 1, 2, 3, 4$, be the involution of $U_n(K,f)$ with eigenspace $R_i^+ = P_i$. Each of u, v, and uv has a positive eigenspace which contains two subspaces P_i; so r_i^* is not one of 1^*, u^*, v^*, or u^*v^*. Now if t is in $c(u,v)$, t must commute with v as $n \neq 6$. If t should anticommute with u, then would $t(P_1) = P_3$, which is impossible as $\dim P_1 \neq \dim P_3$. So t commutes with u and v; hence t leaves invariant P_i, $i = 1, 2, 3, 4$. Consequently t commutes with u and v; thus r_i is in $c(c(u,v))$, which means that $\nu_{u^*} > 4$.

Now let $n = 6$ and u^* be a 2-involution. As above let u be the 2-involution in u^*, and choose P_1, P_2, P_3, and P_4 to be orthogonal nonisotropic subspaces of the eigenspaces U^+ and U^- of u such that $U^+ = P_1 \oplus P_2$, $U^- = P_3 \oplus P_4$ and $\dim P_1 = \dim P_2 = \dim P_3 = 1$. Then $\dim P_4 = 3$. Let v be the 2-involution of $U_n(K,f)$ with eigenspaces $V^+ = P_1 \oplus P_3$ and $V^- = P_2 \oplus P_4$. Clearly $uv = vu$. Define the involutions r_i as above. Then t in $c(u,v)$ must commute with u and v as they are 2-involutions. Hence $t(P_i) = P_i$, $i = 1, 2, 3, 4$; that is, $t(R_i^+) = R_i^+$ and $t(R_i^-) = R_i^-$. Therefore r_i is in $c(c(u,v))$ and $\nu_{u^*} > 4$.

Finally take $n = 6$ and u^* to be a 3-involution. Furthermore, assume E to be not commensurable. Then by Lemma 1.3, there exist nonisotropic 1-dimensional subspaces P_1 and P_4 of U^+ and U^-, respectively, which are not commensurable to each other. Let $P_2 = P_1^\perp \cap U^+$ and $P_3 = P_4^\perp \cap U^-$. Then $E = P_1 \oplus P_2 \oplus P_3 \oplus P_4$ and the 3-involution v with subspaces $V^+ = P_1 \oplus P_3$ and $V^- = P_2 \oplus P_4$ commutes with u. Since $\dim P_1 = \dim P_4 = 1$ and $\dim P_2 = \dim P_3 = 2$, it follows that t in $c(u,v)$ either commutes with both u and v or anticommutes with both of them. However, in the latter case, $t(P_1) = P_4$ in contradiction to the noncommensurability of P_1 and P_4. Thus, t leaves P_i, $i = 1, 2, 3, 4$, invariant, and r_i is in $c(c(u,v))$. So in this case, $\nu_{u^*} > 4$. This proves the theorem.

III. Involutions of the Third Kind.[7]

In this part, u^* will represent an involution of the third kind and u a projective involution contained in u^*. Then $u^2 = \gamma$ in C_n and $\gamma = \lambda^2$ where

---

[7] It is not necessary to determine whether involutions of the second kind are in $\mathfrak{J}^*$. The argument in Part V allows us to pass over this case.

λ is not in Z. Form the group $\Gamma U_n(K_1, f_1)$ acting on the space $E_1 = U^+$ as in Section 1.8. Since u is linear, f_1 will be hermitian in all cases. Let v be a projective involution of $U_n(K, f)$ which commutes or anticommutes with u and let v_1 be the projective semilinear involution of $\Gamma U_n(K_1, f_1)$ which represents v. We first distinguish four cases when v commutes with u and v_1 is in $U_n(K_1, f_1)$ in accordance as to whether v_1 is of the first, second, third, or fourth kinds. Then we treat the case where v_1 is semilinear with respect to the automorphism θ^* described in Section 1.8.

We have already dealt with the case that v_1 is of the first kind in $U_n(K_1, f_1)$ in Section 2.2.

PROPOSITION 3.1. *Let v be a projective involution in $U_n(K, f)$ commuting with u such that its restriction v_1 to E_1 is a projective involution of the first kind in $U_n(K_1, f_1)$. If $n > 3$, then $\nu(u^*, v^*) = 4$.*

3.1. Commuting projective involutions whose restrictions are of the second kind. Assume $n > 4$. Let v_1 be of the second kind in $U_n(K_1, f_1)$. Then $v_1^2 = \beta$ in C_n, $\beta = \mu^2$ with μ in $Z(\lambda)$ and $\mu^T\mu = -1$, where T is the antiautomorphism defined in Section 1.8. Thus $E_1 = V_1^+ \oplus V_1^-$, where V_1^+ and V_1^- are the totally isotropic eigenspaces of v_1, and $\dim_{K_1} V_1^+ = \dim_{K_1} V_1^-$. Let x be in V_1^+ and pick y in V_1^- such that $f_1(x, y) \neq 0$. Then the plane $W_1^+ \subseteq E_1$ generated by x and y is nonisotropic. Set $W_1^- = (W_1^+)^\perp$ and define w_1 to be the 2-involution of $U_n(K_1, f_1)$ with subspaces W_1^+ and W_1^-. As $W_1^+ = xK_1 \oplus yK_1$, v_1 leaves invariant W_1^+ and consequently W_1^-. Thus w_1 commutes with v_1 and its extension w to E is a 2-involution commuting with u and v. The choice of y in V_1^- is not unique. Indeed, let y'' be in $V_1^- \cap W_1^-$ and set $y' = y + y''$. Then $f_1(x, y') = f_1(x, y) \neq 0$, and y and y' are linearly independent. Let w_1' be the 2-involution of $U_n(K_1, f_1)$ with $W_1'^+ = xK_1 \oplus y'K_1$ and $W_1'^- = (W_1'^+)^\perp$ as its eigenspaces. Then as above w_1' commutes with v_1 and its extension w' to E commutes with u and v.

Now let r be in $c(c(u, v))$. As $n > 4$, r commutes with both w and w'. If r anticommutes with u, we form $r' = h_u r$; r' also commutes with w and w'. In any case, the element r_1 of $\Gamma U_n(K_1, f_1)$ which represents r commutes with w_1 and w_1'. Thus $r_1(W_1^+ \cap W_1'^+) = W_1^+ \cap W_1'^+$; that is, $r(x) = x\delta_x$. Furthermore, as r_1 is linear or semilinear with respect to θ^*, we have by Lemma 1.6 that $r_1(x) = x\delta$ for all x in V_1^+, where δ is in K_1. Furthermore, the possibility that r_1 is semilinear relative to θ^* is eliminated as θ^*, not leaving $Z(\lambda)$ fixed, is not an inner automorphism of K_1. So r_1 must be linear and δ is in $Z(\lambda)$. Similarly $r_1(x) = x\delta^-$ for all x in V_1^-, where δ^- is in K_1.

Now $r_1^2 = \epsilon$ in Z implies that $\delta^2 = (\delta^-)^2 = \epsilon$. Hence $\delta^- = \pm\,\delta$; and by Section 1.10, δ is in Z or $\delta = \delta_0\lambda$, where δ_0 is in Z. Now

$$f_1(x,y) = f_1(r_1(x), r_1(y)) = \delta^T\delta^- f_1(x,y)$$

for x in V_1^+ and y in V_1^-. As we can further stipulate that $f_1(x,y) \neq 0$, we have that $\delta^T\delta = \pm 1$ according as $\delta^- = \pm\,\delta$.

If δ is in Z and $\delta^- = \delta$, then $\delta^T = \delta^J$ and δ is in C. If δ is in Z and $\delta^- = -\delta$, by replacing r_1 by v_1r_1 in the above argument we obtain that $v_1r_1(x) = x\delta\mu$ for x in V_1^+ and V_1^- and that $\delta\mu$ is in C. If $\delta = \delta_0\lambda$, we replace r_1 by u_1r_1 or $u_1v_1r_1$ to obtain, respectively, that $u_1r_1(x) = x\delta\lambda$ or $u_1v_1r_1(x) = x\delta\mu\lambda$ for x in V_1^+ and V_1^-, where $\delta\lambda$ or $\delta\mu\lambda$ is in C. In any case, we see that there exists an η in C such that ηr_1 coincides with the restriction of 1, u, v, or uv to E_1. Hence r^* is one of 1^*, u^*, v^*, or u^*v^*. Thus we have shown:

PROPOSITION 3.2. *Let v be a projective involution in $U_n(K,f)$ commuting with u whose restriction v_1 to E_1 is a projective involution of the second kind in $U_n(K_1,f_1)$. If $n > 4$, $\nu(u^*,v^*) = 4$.*

3.2. Commuting projective involutions whose restrictions are of the third kind. Assume $n > 2$. Let v be a projective involution of $U_n(K,f)$ commuting with u whose restriction v_1 to E_1 is of the third kind in $U_n(K_1,f_1)$. Then $v_1^2 = \beta$ is in C_n and $\beta = \mu^2$, where μ is not in $Z(\lambda)$ but is in K_1. Let K_2 be the fixed subsfield of K_1 by μ^*; then by Section 1.8, $K_1 = K_2 \oplus K_2\psi$, where $\psi^{\mu^*} = -\psi$ and ψ^2 is in $Z(\lambda)$. As in Section 1.8, consider E_1 to be a vector space E_1^* over K_2. Then $E_1^* = V_1^+ \oplus V_1^-$, where V_1^+ and V_1^- are the eigenspaces of v_1. We form the group $\Gamma U_n(K_2,f_2)$ acting on $E_2 = V_1^+$. Since v_1 is linear in $U_n(K_1,f_1)$, f_2 is always hermitian.

Let K_1' be the fixed subsfield of K by $(\mu^{-1})^* = \mu^*$. Then $K = K_1' \oplus K_1'\psi$. By Section 1.8, the center of K_1' is $Z(\mu)$. As $\mu^{\lambda^*} = \mu$, $\lambda^{\mu^*} = \lambda$; and λ is in K_1'. But $Z(\mu) \neq Z(\lambda)$ as they are both quadratic extensions of Z and μ is not in $Z(\lambda)$. So λ is not in $Z(\mu)$. Hence by Section 1.8, $K_1' = K_2' \oplus K_2'\chi$, where χ is an element of K_1' such that χ^2 is in $Z(\mu)$ and $\chi^{\lambda^*} = -\chi$. Noting that $K_1\theta = K_1\chi$, we shall replace θ by χ. Then we have the following relations:

(i) $K = K_2 \oplus K_2\psi \oplus K_2\chi \oplus K_2\psi\chi$.

(ii) $\lambda^{\chi^*} = -\lambda$; $\mu^{\chi^*} = \mu$; $\lambda^{\psi^*} = \lambda$; $\mu^{\psi^*} = -\mu$.

(iii) χ^2 is in $Z(\mu)$; ψ^2 is in $Z(\lambda)$.

Since v_1 is of the third kind in $U_n(K_1,f_1)$, we have by Section 1.8 that the center of K_2 is $Z(\lambda,\mu)$.

Now let t be in $c(u,v)$. Then if t commutes with u, t is represented in $U_n(K_1,f_1)$ by its restriction to E_1. In accordance with whether t commutes or anticommutes with v, t_1 is represented by a projective semilinear involution t_2 in $\Gamma U_n(K_2,f_2)$ which is linear or semilinear relative to the automorphism ψ^* of K_2. Furthermore, $t^2=\eta$ with η in $Z(\mu)$. If t anticommutes with u, then t is represented in $\Gamma U_n(K_1,f_1)$ by a projective semilinear involution t_1 which is relative to the automorphism χ^* of K_1. In this case, $t_1{}^2=\eta_1$, where η_1 is in $Z(\mu)$. Then, as $\mu^{\chi^*}=\mu$, t_1 in turn is represented in $\Gamma U_n(K_2,f_2)$ by a projective semilinear involution t_2 which is relative to the automorphism χ^* or $\chi^*\psi^*$ according as t_1 commutes or anticommutes with v_1. Finally $t_2{}^2=\eta_2$, where η_2 is in $Z(\lambda,\mu)$.

The extension w of a 1-involution w_2 of $U_n(K_2,f_2)$ to E is a 1-involution in $c(u,v)$. Consequently, a projective involution r of $c(c(u,v))$ commutes with w as $n>2$. Therefore, if r_2 represents r in $\Gamma U_n(K_2,f_2)$, r_2 commutes with w_2. This means that r_2 leaves every 1-dimensional nonisotropic space of E_2 invariant. Furthermore, r_2 is a projective semilinear involution relative to an involutive automorphism of K_2 and $r_2{}^2=\epsilon$ with ϵ in $Z(\lambda,\mu)$. So by Lemma 1.5, $r_2(x)=x\delta$ with δ in K_2 for all x in E_2. Moreover, r_2 must be semilinear with respect to an inner automorphism of K_2. Since χ^*, ψ^* and $\chi^*\psi^*$ are not inner automorphisms, we have that r_2 must be linear. Therefore, r commutes with u and v; so r_2 is the restriction of r to E_2 and ϵ is in Z. Furthermore, δ is in $Z(\lambda,\mu)$ and $\delta^2=\epsilon$. We now apply the result in Section 1.10 twice. First either δ is in $Z(\lambda)$ or $\delta=\delta_1\mu$ with δ_1 in $Z(\lambda)$. Then the square of δ or δ_1, respectively, is an element in Z. This implies that δ is Z, $\delta=\delta_1\mu$, $\delta=\delta_2\lambda$, or $\delta=\delta_3\lambda\mu$ with δ_i in Z. Let u_2 and v_2 be the restrictions of u and v to E_2. Let r_2' be that one of r_2, $u_2{}^{-1}r_2$, $v_2{}^{-1}r_2$, or $u_2{}^{-1}v_2{}^{-1}r_2$ such that $r_2'(x)=x\delta_i$ for x in E_2 and some $i=1,2,3,4$. Since r_2 is in $U_n(K_2,f_2)$, δ is in Z, and E_2 is nonisotropic, we have $\delta^T\delta=\delta^J\delta=1$. Therefore, the extension r' of r_2' to E is in C_n and r is in one of the cosets 1^*, u^*, v^* or u^*v^*. This gives us the following result.

Proposition 3.3. *Let v be a projective involution in $U_n(K,f)$ commuting with u such that its restriction v_1 to E_1 is a projective involution of the third kind in $U_n(K_1,f_1)$. If $n>2$, $\nu(u^*,v^*)=4$.*

3.3. Commuting projective involutions whose restrictions are of the fourth kind.

Let v_1 be of the fourth kind in $U_n(K_1,f_1)$. Assume $n>4$. Then $v_1{}^2=\beta$ with β in C and β not a square in K_1. As in Section 1.9, extend the sfield K_1 to a sfield L_1 by adjoining a square root ρ of β. The center L_1 is then $Z(\lambda, \rho)$. Then E_1 can be considered as a right vector space F_1 over L_1 by setting $x\rho = v_1(x)$ and f_1 can be extended to a nondegenerate hermitian sesquilinear form g_1 defined on F_1. We then form the group $\Gamma U_m(L_1, g_1)$ acting on F_1.

Let t_1 be the restriction of a projective involution in $U_n(K, f)$ which commutes with u or the representation of a projective involution which anticommutes with u, in which case t_1 is semilinear with respect to the automorphism θ^* of K_1. If, furthermore, t commutes or anticommutes with v, t_1 commutes or anticommutes, respectively, with v_1. Then t_1 can be considered as a projective semilinear involution in $\Gamma U_m(L_1, g_1)$ relative to one of the automorphisms 1, θ^*, σ or $\theta^*\sigma$ where σ is the automorphism of L_1 leaving K_1 fixed and defined by $\sigma : \rho \rightarrow -\rho$. It is easily verified that σ and θ^* commute; so all these automorphisms are involutive. The projective involutions u, v, and uv are represented in $U_m(L_1, g_1)$ by the homothetic transformations h_λ, h_ρ, and $h_{\lambda\rho}$, respectively.

Every 1-involution w_1 of $U_m(L_1, g_1)$ extends to a 2-involution w of $U_n(K, f)$ in $c(u, v)$. *If* r_1 in $\Gamma U_m(L_1, g_1)$ represents a projective involution r in $c(c(u, v))$, then r must commute with w as $n > 4$ and r_1 must commute with w_1. So r_1 leaves invariant every 1-dimensional nonisotropic subspace of F_1. Furthermore, r_1 is a projective semilinear transformation relative to an involutive automorphism of L_1 and $r_1^2 = \epsilon$ with ϵ in $Z(\lambda)$. So by Lemma 1.5, $r_1(x) = x\delta$ with δ in L_1 for all x in F_1. However, r_1 cannot be semilinear with respect to the automorphisms, θ^*, σ, or $\theta^*\sigma$ as none of these are inner automorphisms. Consequently, r_1 is linear and must be the representation of a projective involution r commuting with u and v. Therefore, ϵ is in Z and δ is in $Z(\lambda, \rho)$. Also $\delta^2 = \epsilon$. So applying the result of Section 1.10 twice as in Section 3.2, we have that $\delta = \delta_0$, $\delta = \delta_1\rho$, $\delta = \delta_2\lambda$ or $\delta = \delta_3\lambda\rho$, where δ_i is in Z and $\delta_i^J\delta_i = 1$. Thus r_1 is the representation of a projective involution in 1^*, u^*, v^*, or u^*v^*. This yields the following result.

PROPOSITION 3.4. *Let v be a projective involution in $U_n(K, f)$ commuting with u such that its restriction v_1 to E_1 is a projective involution of the fourth kind in $U_n(K_1, f_1)$. If $n > 4$, $\nu(u^*, v^*) = 4$.*

Now we consider the case that v anticommutes with u. Then as in Section 1.8, we set $v' = h_\theta v$; v' is a projective semilinear involution in $\Gamma U_n(K, f)$ which commutes with $u' = \lambda^{-1}u$. The restriction v_1 of v' to E_1 is a projective semilinear involution in $\Gamma U_n(K_1, f_1)$ relative to the automorphism θ^* of K_1, and $v_1^2 = \beta$ with β in Z. First note that if $v_1(x) = x\mu$ for any x in E_1, $v_1^2(x) = x\mu\mu^{\theta^*} = x\beta$. Therefore, we distinguish two cases:

Case I, where $\beta = \mu\mu^{\theta^*}$ for some μ in K_1, and Case II, where β is not of the form $\mu\mu^{\theta^*}$ for any μ in K_1.

3. 4. Anticommuting projective involutions. Case I. Assume $n > 2$. Let v be a projective involution in $U_n(K, f)$ anticommuting with u such that v is represented by a projective semilinear involution v_1 in $\Gamma U_n(K_1, f_1)$, where $v_1{}^2 = \beta$, and $\beta = \mu\mu^{\theta^*}$ is in Z with μ in K_1. Let K_2 be the fixed subsfield of K_1 by $\theta^*\mu^*$. Since μ is in K_1, $\lambda^{\mu^*} = \lambda$ and $\lambda^{\theta^*\mu^*} = -\lambda$; so $K_1 = K_2 \oplus K_2\lambda$. Let Z_2 be the center of K_2; then an element in Z_2 commutes not only with every element of K_2 but also with λ and therefore with every element of K_1. So $Z_2 \subseteq K_2 \cap Z(\lambda) = Z$. Since Z is left elementwise invariant by $\theta^*\mu^*$, $Z \subseteq Z_2$. Therefore $Z = Z_2$. In the manner of Section 1. 8, form the group $\Gamma U_n(K_2, f_2)$ acting on $E_2 = V^+$. The form f_2 will be hermitian or alternating.

Now an element t of $c(u, v)$ which commutes with u and v commutes also with $h_\theta v$ and so can be represented in $U_n(K_2, f_2)$ or $Sp_n(K_2, f_2)$ by the restriction t_2 of t to E_2. If an element t commutes with u and anticommutes with v, then ut will commute with both u and $h_\theta v$ and hence ut can be restricted to a linear transformation in $U_n(K_2, f_2)$ or $Sp_n(K_2, f_2)$. Similarly, if t anticommutes with u and commutes with v, form vt; if t anticommutes with both u and v, form uvt.

In the case f_2 is hermitian, for every 1-dimensional subspace of E_2 which is nonisotropic relative to the form f_2, there exists a 1-involution w_2 in $U_n(K_2, f_2)$ which can be extended to a 1-involution w of $U_n(K, f)$ that commutes with u and v. Denote by r' that one of r, ur, vr, or uvr which commutes with u and v. Then as $n > 2$, r' will commute with the 1-involutions w in $c(u, v)$, and its restriction r_2 to E_2 will commute with the 1-involutions w_2 of $U_n(K_2, f_2)$. Hence by Lemma 1. 5, $r_2(x) = x\delta$ with δ in Z for every vector x in E_2. Furthermore $\delta^H\delta = 1$, where H is the antiautomorphism of K_2 corresponding to f_2. As δ is in Z, $\delta^H = \delta^J$. So the extension r' of r_2 to E is a homothetic transformation of E in $U_n(K, f)$. Therefore, r is in one of 1^*, u^*, v^* or u^*v^* and $\nu(u^*, v^*) = 4$.

In the case f_2 is alternating, any plane in E_2 generated by two nonorthogonal vectors will be nonisotropic. So if x is any vector in E_2, pick two linearly independent vectors y and y' not in $(xK)^\perp$ and call W and W' the nonisotropic planes generated by x and y, and x and y', respectively. Then the 2-involutions w_2 and w_2' with W and W' as their respective positive eigenspaces can be extended to 2-involutions w and w' in $c(u, v)$. If r' is the projective involution chosen in the previous paragraph, r' commutes with both w and w'

when $n > 4$ while w and w' are 2-involutions. Then the restriction r_2 of r' to E_2 commutes with w_2 and w_2'; so $r_2(W \cap W') = W \cap W'$. That is, $r_2(x) = x\delta_x$ for each x in E_2. Then by Lemma 1.5, $r_2(x) = x\delta$ with δ in Z for all x in E_2. As it was argued in the above paragraph, r is in one of 1^*, u^*, v^*, or u^*v^*.

PROPOSITION 3.5. *Let v be a projective involution of $U_n(K,f)$ anticommuting with u such that v is represented by a projective semilinear involution v_1 of $\Gamma U_n(K_1,f_1)$, where $v_1^2 = \beta$ and $\beta = \mu\mu^{\theta^*}$. If $n > 2$ when f_2 is herimitian and $n > 4$ when f_2 is alternating, $\nu(u^*,v^*) = 4$.*

3.5. Anticommuting projective involutions. Case II. Assume $n > 4$. Now consider the case that v anticommutes with u and is represented by a projective semilinear involution v_1 acting on E_1 such that $v_1^2 = \beta$, where β is not of the form $\mu\mu^{\theta^*}$ for μ in K_1. We then adjoin a square root ρ of β to K_1 forming a sfield $L_1 = K_1(\rho)$ and form the group $\Gamma U_m(L_1,g_1)$ acting on the space F_1 as in Section 1.9. Now ρ satisfies the relation $\phi\rho = \rho\phi^{\theta^*}$ for all ϕ in K_1. In particular if ϕ is in Z_1, the center of L_1, then $\phi^{\theta^*} = \rho^{-1}\phi\rho = \phi$. On the other hand, $\phi(\xi + \rho\eta) = (\xi + \rho\eta)\phi$ for all ξ and η in K_1. This implies that $\phi\xi - \xi\phi = \rho(\eta\phi - \phi\eta)$. So ϕ must commute with ξ and η; hence $Z_1 \subseteq Z(\lambda)$. This implies that $Z_1 = Z$.

Again an element t in $U_n(K,f)$ which commutes with u and v commutes with $h_\theta v$ and hence can be restricted to a projective involution t_1 in $U_n(K_1,f_1)$ which commutes with v_1. Hence t_1 can be considered to be in $U_m(L_1,g_1)$. If t in $c(u,v)$ anticommutes with one or both of u, v, or uv, then one of ut, vt, or uvt will commute with both u and v.

Now let w_1 be an 1-involution of $U_m(L_1,g_1)$. Then w_1 can be extended to a 2-involution w in $U_n(K,f)$ which commutes with u and v. Since $n > 4$, r in $c(c(u,v))$ must commute with w. Denote by r' that one of r, ur, vr, or uvr which commutes with u and v. Then r' commutes with w and is represented by a linear projective involution r_1' of $\Gamma U_m(L_1,g_1)$ commuting with w_1. Hence r_1' leaves invariant every 1-dimensional nonisotropic subspace of F_1. Therefore, by Lemma 1.5 $r_1'(x) = x\delta$ with δ in Z for all x in F_1. Again as in Section 3.4, this implies that $\delta^J\delta = 1$ and thus that r' is in 1^*. Hence r is in one of 1^*, u^*, v^*, or u^*v^*. This yields the following result.

PROPOSITION 3.6. *Let v be a projective involution of $U_n(K,f)$ anticommuting with u such that v is represented by a projective semilinear involution v_1 of $U_n(K_1,f_1)$, where $v_1^2 = \beta$ and β is not of the form $\mu\mu^{\theta^*}$ for μ in K_1. If $n > 4$, $\nu(u^*,v^*) = 4$.*

Now Propositions 3.1 through 3.6 yield the following theorem.

THEOREM 3.1. *If the dimension of E is greater than* 4, *and if u^* is an involution of the third kind in $PU_n(K,f)$, then $\nu_{u^*}=4$.*

IV. Involutions of the Fourth Kind.

In this part, u^* will represent an involution of the fourth kind in $PU_n(K,f)$. We shall assume $n>4$. Pick a representative u in u^*; then $u^2=\gamma$ in C_n, where γ is not a square in K. As in Section 1.9, form the sfield $L=K(\rho)$, where $\rho^2=\gamma$. The center of L is $Z(\rho)$. We consider E to be a right vector space F over L by setting $x\rho=u(x)$, and we form the group $\Gamma U_m(L,g)$. Let σ be the automorphism of L leaving K fixed such that $\rho^\sigma=-\rho$. Then an element of $U_n(K,f)$ which commutes or anticommutes with u can be considered to be an element of $\Gamma U_n(K,f)$ which is, respectively, linear or semilinear relative to the automorphism σ.

We shall designate by v a projective involution commuting or anticommuting with u. We distinguish various cases and in each show that $\nu(u^*,v^*)=4$.

4.1. Projective involutions of the first kind in $U_m(L,g)$. Let v be a projective involution of the first kind in $U_m(L,g)$. Then $v^2=\beta$ with β in C, $\beta=\mu^2$, where μ is in $Z(\rho)$ and $\mu^J\mu=1$. As in Section 1.6, $F=V^+\oplus V^-$, where V^+ and V^- are the orthogonal nonisotropic eigenspaces. Since $m>2$, one of V^+ and V^-, say V^+, must be of L-dimension greater than 1. Let x be a nonisotropic vector in V^+ and w the 1-involution of $U_m(L,g)$ with xK as its 1-dimensional subspace W^+. Since $W^+\subseteq V^+$ and v is in $U_m(L,g)$, v leaves invariant W^+ and $W^-=(W^+)^\perp$. Consequently, w commutes with v, and thus, w is in $c(u,v)$. Let r be in $c(c(u,v))$. Then as an element of $\Gamma U_m(L,g)$, r is linear or semilinear with respect to the automorphism σ. Furthermore, r must commute with w as $m>2$; hence $r(x)=x\delta_x$ with δ_x in L for each nonisotropic vector x in V^+. So by Lemma 1.5, $r(x)=x\delta$ with δ in L for x in V^+. Also, r must be semilinear with respect to an inner automorphism of L. But σ does not leave fixed the center $Z(\rho)$ of L and, therefore, is not an inner automorphism of L. So r is linear and δ is in $Z(\rho)$. Also $r^2=\epsilon$ with ϵ in Z; so $\delta^2=\epsilon$. It then follows from Section 1.10 that δ is in Z or $\delta=\delta_0\rho$ with δ_0 in Z.

If $\dim_L V^->1$, it follows in a similar manner that $r(x)=x\delta^-$ with δ^- in L for x in V^-; and as $r^2=\epsilon$, we have $(\delta^-)^2=\delta^2=\epsilon$. So $\delta^-=\pm\delta$. If $\dim_L V^-=1$, then $r(x)=x\delta_x$ for x in V^-. Again, $\delta_x{}^2=\delta^2$; so $\delta_x=\pm\delta$.

So δ_x is in $Z(\rho)$ and hence is independent of x in V^-. Set $\delta_x = \delta^-$. Since g is sesquilinear with respect to the antiautomorphism J extended to L, and V^+ is nonisotropic, we have that $\delta^J\delta = 1$. Thus if δ is in Z, δ is in C. If $\delta = \delta_0\rho$, $\delta_0{}^J\delta_0 = 1$ as $\rho^J = \rho^{-1}$; so δ_0 is in C. In any case r considered an element of $U_n(K,f)$ is in one of the cosets 1^*, u^*, v^*, or u^*v^*. So we have proved the following result.

PROPOSITION 4.1. *Let v be a projective involution of the first kind in $U_m(L,g)$. If $n > 4$, $\nu(u^*,v^*) = 4$.*

4.2. Projective involutions of the second kind in $U_m(L, g)$. Let v be a projective involution of the second kind in $U_m(L,g)$. Then $v^2 = \beta$ with β in C, $\beta = \mu^2$ with μ in $Z(\rho)$, and $\mu^J\mu = -1$. As in Section 1.7, $F = V^+ \oplus V^-$, where V^+ and V^- are totally isotropic subspaces of L-dimension $m/2$.

Let x be an arbitrary vector in V^+ and pick y in V^- such that $g(x,y) \neq 0$. Then $W^+ = xL \oplus yL$ is nonisotropic and the 2-involution w in $U_m(L,g)$ with W^+ and $W^- = (W^+)^\perp$ as its eigenspaces commutes with v. Also $W^- = W^- \cap V^+ \oplus W^- \cap V^-$, where $W^- \cap V^+$ and $W^- \cap V^-$ are totally isotropic subspaces of L-dimension $m/2 - 1$. Since $m > 2$, these subspaces are not zero; so there exists a vector b in $W^- \cap V^-$. Set $y' = y + b$; then $g(x,y') = g(x,y) \neq 0$. Let w' be the 2-involution of $U_m(L,g)$ with $W'^+ = xL \oplus y'L$ and $W'^- = (W')^\perp$ as its eigenspaces. Both w and w' can be considered to be in $c(u,v)$ when considered as transformations of $U_n(K,f)$.

If $n \neq 8$, we can conclude that r in $c(c(u,v))$ commutes with w and w'. We shall show that this is still true when $n = 8$. Then

$$\dim_L W^\pm \cap V^\pm = \dim_L W'^\pm \cap V^\pm = 1.$$

We first assume that r commutes with w; then we show that r commutes with w'. Indeed, if r anticommutes with w', then $r(x)$ is not in $W^+ \cap V^+ = W^+ \cap W'^+$ and so is in $W^+ \cap V^- = yL$ as r must also commute or anticommute with v. Thus $r(x) = y\phi$ with ϕ in L. But also $r(x)$ is in W'^- and so is orthogonal to x. This is a contradiction as $g(x,y) \neq 0$; hence r commutes with w' in this case.

A similar argument shows that if r commutes with w', r commutes with w. So there remains to consider only the case that r anticommutes with w and w'. We shall exclude this case by showing that there exists a third involution w'' in $c(u,v)$ with which r commutes. Then the argument of the previous paragraph will yield an obvious contradiction.

Inasmuch as b is orthogonal to x, y, and y', we have that b is in $W^- \cap W'^- \cap V^-$. It then follows that $r(x) = b\psi$, where ψ is in L. We replace $b\psi$ by b so that $r(x) = b$. Because x and $r(x)$ are in V^+ and V^-, respectively, r anticommutes with v. Thus it follows that there exists a vector a in $V^+ \cap W^-$ such that $r(a) = y$. Set $z = x + a$; then $r(z) = y'$. Also if $\epsilon = r^2$, $r(y) = a\epsilon$ and $r(b) = x\epsilon$.

Of course, it may result that z and $r(z)$ are orthogonal. However, we shall now show that there exists a vector $z_\phi = x + a\phi$ with ϕ in L such that $g(z_\phi, r(z_\phi)) \neq 0$. If τ is the automorphism related to r, then $r(z_\phi) = y\phi^\tau + b$. Setting $\xi = g(x,y) \neq 0$, we have

$$g(z_\phi, r(z_\phi)) = \xi\phi^\tau + \phi^J\xi^J\epsilon. \tag{4.1}$$

If $g(z_\phi, r(z_\phi)) = 0$ for all ϕ in L, we have

$$\phi^J = -\xi\phi^\tau\xi^{-J}\epsilon^{-1}. \tag{4.2}$$

Setting $\phi = 1$ in (4.2), we obtain $\xi^{-J}\epsilon^{-1} = -\xi^{-1}$ and

$$\phi^J = \xi\phi^\tau\xi^{-1} = \phi^{\tau\xi^*}. \tag{4.3}$$

Thus L must be commutative as J is an antiautomorphism. Then it follows from (4.3) that $J = \tau$. There are two possibilities for τ: either $\tau = 1$ or $\tau = \sigma$. In either case, J is the identity on K. Consequently $U_8(K,f)$ is the orthogonal group $O_8(K,f)$. Then by Section 1.9, $\rho = i = \sqrt{-1}$ and $i^J = -i$. Thus $J = \sigma$. But β is in C; so $\beta^J\beta = \beta^2 = 1$. It is easily seen that if $\beta = 1$, then $\mu^J\mu = 1$. If $\beta = -1$, then $\mu = \pm i$; and again $\mu^J\mu = 1$. However, the condition that v is a projective involution of the second kind in $U_m(L,g)$ requires that $\mu^J\mu = -1$. Thus we can conclude that there exists an element ϕ in L such that $g(z_\phi, r(z_\phi)) \neq 0$.

Clearly then the subspace $W''^+ = z_\phi L \oplus r(z_\phi)L$ is nonisotropic. So the 4-involution w'' of $U_8(K,f)$ with W''^+ as its positive eigenspace is in $c(u,v)$. Clearly $r(W''^+) = W''^+$; so r must commute with w''. Therefore, w'' is the involution which we seek.

Consequently when $n > 4$, $r(W^+ \cap W'^+) = W^+ \cap W'^+$; that is, $r(x) = x\delta_x$ for all x in V^+. Then by Lemma 1.6, $r(x) = x\delta$ with δ in L for all x in V^+. Furthermore, r cannot be semilinear relative to σ as σ is not an inner automorphism of L. Therefore, r is linear and δ is in $Z(\rho)$. Furthermore, $\delta^2 = \epsilon$ is in Z, where $\epsilon = r^2$. So by Section 1.10, δ is in Z or $\delta = \delta_0\rho$, where δ_0 is in Z. A similar argument shows that $r(x) = x\delta^-$ with δ^- in $Z(\rho)$ for all x in V^-. We then have $\delta^- = \pm\delta$.

If δ is in Z and $\delta^- = \delta$, r is the homothetic mapping h_δ; and as F is nonisotropic, $\delta'\delta = 1$. Thus $r^* = 1^*$ in this case. If δ is in Z and $\delta^- = -\delta$, then vr is the homothetic mapping $h_{\mu\delta}$. So in this case $r^* = v^*$. If $\delta = \delta_0\rho$ and $\delta^- = \delta$, then similarly $r^* = u^*$. Finally, if $\delta = \delta_0\rho$ and $\delta^- = -\delta$, then $r^* = u^*v^*$. Thus we have proved:

PROPOSITION 4.2. *Let v be a projective involution of the second kind in $U_m(L, g)$. If $n > 4$, $\nu(u^*, v^*) = 4$.*

4.3. Projective involutions of the third kind in $U_m(L, g)$. Let v be a projective involution of the third kind in $U_m(L, g)$. Then $v^2 = \beta$, $\beta = \mu^2$, where μ is in L but not in $Z(\rho)$. As in Section 1.8, let L_1 be the fixed subfield of L by μ^* and let θ be an element of L such that $\theta^{\mu^*} = -\theta$ and θ^2 is in $Z(\rho)$. Then $L = L_1 \oplus L_1\theta$. The center of L_1 is $Z(\rho, \mu)$. F can be considered as a right vector space F^* over L_1, and $F^* = V^+ \oplus V^-$. Set $F_1 = V^+$, and take the form g_1 as in Section 1.8. Then form the group $\Gamma U_m(L_1, g_1)$ acting on the space F_1. As v is linear, g_1 is hermitian.

If t is a projective involution of $U_m(L, g)$ which commutes with v, then by Section 1.8, t can be restricted to a projective involution t_1 of $U_m(L_1, g_1)$. Conversely, if t_1 is a projective involution of $U_m(L_1, g_1)$, t_1 can be extended to a projective involution t of $U_m(L, g)$. If t anticommutes with v, then t is represented by a projective semilinear involution t_1 of $\Gamma U_m(L_1, g_1)$ which is semilinear relative to the automorphism θ^* of L_1.

The situation is not quite so simple if t is a projective semilinear involution of $\Gamma U_m(L, g)$ relative to the automorphism σ. If we can show that $\mu^\sigma = \mu$, then by Section 1.8 we can restrict t to a semilinear projective involution t_1 of $\Gamma U_m(L_1, g_1)$ which is semilinear with respect to the automorphism σ. Indeed, as we shall see, this is true if t is in $c(c(u, v))$ and t commutes with v. Furthermore, if t is such a projective semilinear involution in $c(c(u, v))$ which anticommutes with v, then $\mu^\sigma = -\mu$; but then we show that t has a restriction t_1 to F_1 which is a projective semilinear involution of $\Gamma U_m(L_1, g_1)$ relative to the automorphism σ of L.

Therefore, let t be a projective semilinear involution of $\Gamma U_m(L, g)$ relative to the automorphism σ of L which is in $c(c(u, v))$. Let x be a vector of F_1 which is nonisotropic with respect to the form g_1. Set $W_1^+ = xL_1$ and $W_1^- = (W_1^+)^\perp$, and let w_1 be the 1-involution of $U_m(L_1, g_1)$ with W_1^+ and W_1^- as its eigenspaces. As $m > 2$, a projective semilinear involution r in $c(c(u, v))$ must commute with the extension w of w_1 to F. In particular, t commutes with w. As a transformation of F, w has positive eigenspace $W^+ = xL_1 \oplus xL_1\theta = xL$. Therefore $t(x) = x\delta_x$ with δ_x in L. Set $\delta_x = \delta$.

Since $t^2 = \epsilon$ in Z, we have $\delta\delta^\sigma = \epsilon$. Then it is easily verified that $\delta^*\sigma = \sigma(\delta^\sigma)^*$; so the automorphism $\pi = \delta^*\sigma$ is an involution. Let L_0 be the fixed subsfield of L by π. Then L_0 is of index 2 in L. Furthermore, $\rho^\pi = \rho^{\delta^*\sigma} = \rho^\sigma = -\rho$. So $L = L_0 \oplus L_0\rho$.

We next show that $L_0 = K$. Let ϕ be in K; then $\phi = \phi_1 + \rho\phi_2$, where ϕ_1 and ϕ_2 are in L_0. But $\phi^\sigma = \phi$; so $\phi_1^\sigma - \rho\phi_2^\sigma = \phi_1 + \rho\phi_2$. This implies that $\phi_1^\sigma = \phi_1$ and $\phi_2^\sigma = -\phi_2$. So ϕ_1 is in K and $\phi_2 = \rho\phi_3$ with ϕ_3 in K. Thus $\phi = \phi_1 + \phi_3\gamma = \phi_1 + \rho\phi_2$, and the representation of ϕ is unique only if $\phi_2 = 0$ and $L_0 = K$.

But then the restriction of π to $L_0 = K$ coincides with the restriction of δ^* to L_0. But also π is the identity on L_0; so δ commutes with every element of K. Also δ will commute with every element of $K\rho$; so δ is in $Z(\rho)$. Now $t(v(x)) = t(x)\mu^\sigma = x\delta\mu^\sigma$ and $v(t(x)) = v(x)\delta = x\mu\delta$. So if t commutes with v, $\mu^\sigma = \mu$ and the method of Section 1.8 applies and we can restrict t to F_1.

If t anticommutes with v, then $\mu^\sigma = -\mu$. Therefore, $\mu^*\sigma = \sigma\mu^*$ and σ leaves L_1 invariant. Furthermore, t also anticommutes with $h_{\mu^{-1}}$; so t commutes with $v' = h_{\mu^{-1}}v$ and consequently can be restricted to a semilinear projective involution of $\Gamma U_m(L_1, g_1)$ acting on F_1 relative to the automorphism σ of L_1.

So in general, a projective semilinear involution r in $c(c(u,v))$ can be restricted to or is represented by a projective semilinear involution r_1 in $\Gamma U_m(L_1, g_1)$ which is linear or semilinear with respect to the involutive automorphism θ^* or σ of L_1. Since the choice of a nonisotropic vector x in F_1 was arbitrary, r_1 must commute with every 1-involution in $U_m(L_1, g_1)$. Hence r_1 leaves invariant every 1-dimensional nonisotropic subspace of F_1. Since $r_1^2 = \epsilon$ with ϵ in $Z(\rho, \mu)$, Lemma 1.5 yields that $r(x) = x\delta$ with δ in L_1 for all x in F_1. Furthermore, r_1 must be semilinear with respect to an inner automorphism of L_1. However, neither θ^* nor σ is an inner automorphism. Consequently, r_1 must be linear; δ is in $Z(\rho, \mu)$; and $\delta^2 = \epsilon$ with ϵ in Z. Then, as it was argued in Section 3.3, r is in one of 1^*, u^*, v^*, or u^*v^*. This yields the following result.

PROPOSITION 4.3. *Let v be a projective involution of the third kind in $U_m(L, g)$. If $n > 4$, $\nu(u^*, v^*) = 4$.*

4.4. Projective involutions of the fourth kind in $U_m(L, g)$.

Let v be a projective involution of the fourth kind in $U_m(L, g)$. Then $v^2 = \beta$ in C_n and β is not a square in L. We adjoin a square root ρ' of β to L forming the sfield $M = L(\rho') = K(\rho, \rho')$. The center of L is $Z(\rho, \rho')$. As in Section 1.9, the space F can be considered as a right vector space G over M of dimension $q = m/2 = n/4$. The form g can be extended to a nondegenerate hermitian sesquilinear form h defined on G relative to the extension of the antimorphism J to M. We then form the group $\Gamma U_q(M, h)$. An element in $U_n(K, f)$ which commutes with u and v can be considered to be in $U_q(M, h)$ and conversely. An element in $U_n(K, f)$ which anticommutes with u and commutes with v, or which commutes with u and anticommutes with v, or which anticommutes with both u and v can be considered to be in $\Gamma U_q(M, h)$ relative to the automorphisms σ, σ', or $\sigma\sigma'$, respectively, where $\sigma : \rho \rightarrow -\rho$, $\sigma : \rho' \rightarrow \rho'$, $\sigma' : \rho' \rightarrow -\rho'$, $\sigma' : \rho \rightarrow \rho$, and σ and σ' leave K fixed.

Now let w be a 1-involution in $U_q(M, h)$; w can be considered as a 4-involution in $U_n(K, f)$ which commutes with both u and v. If r is in $c(c(u, v))$ then r commutes or anticommutes with w. If $n > 8$, we can conclude immediately that r commutes with w. Let $n = 8$ and assume r anticommutes with w; then r, as an element of $\Gamma U_q(M, h)$, is such that $r(x) = y$, where x is in W^+, the positive eigenspace of w, and where y is in W^-, the negative eigenspace of w. The vector y is orthogonal to x and is nonisotropic as $h(y, y) = h(x, x) \neq 0$. Also $r^2 = \epsilon$ with ϵ in Z, so $r(y) = r^2(x) = x\epsilon$. By Lemma 1.4, for every element ϕ of M, there exists an element $\psi = \phi$, $\phi + 1$, or $\phi - 1$ such that $z_\psi = x + y\psi$ is nonisotropic; and r commutes or anticommutes with the 1-involution w_ψ in $U_n(M, h)$ with $z_\psi M$ as its positive eigenspace. In the first case $r(z_\psi) = z_\psi \delta$. This yields $y + x\psi^\tau \epsilon = x\delta_\psi + y\psi\delta_\psi$ where $\tau = 1, \sigma, \sigma'$, or $\sigma\sigma'$ is the automorphism of M associated with r. Due to the linear independence of x and y, we obtain first that $\delta_\psi = \psi^{-1}$ and then that $\psi^\tau \psi = \epsilon^{-1}$, which is in Z. If r anticommutes with w_ψ, then $r(z_\psi)$ is orthogonal to z_ψ. Hence $h(z_\psi, r(z_\psi)) = \xi\psi^\tau\epsilon + \psi^J \xi = 0$ where $\xi = h(x, x) = h(y, y) \neq 0$. This yields $\psi^J = -\xi\psi^\tau\xi^{-1}\epsilon$.

Now by the proper choice of ϕ, we shall obtain contradictions to two of the equations of the above paragraph and thus conclude that r commutes with w. We first note that $Z \neq GF(3)$; otherwise $\beta = \pm 1$ and $\gamma = \pm 1$. This is impossible inasmuch as neither β nor γ can be equal to 1 without u and v being of the first kind in $U_n(K, f)$ while if $\beta = \gamma = -1$, then v would be of the first kind in $U_m(L, g)$. So we choose $\omega_0 \neq 0, \pm 1$ in Z and take $\phi = \rho + \rho' + \omega_0$. Then $\psi = \rho + \rho' + \omega$ where $\omega \neq 0$ and is in Z. Computation shows that

$$(4.4) \quad \psi^\tau \psi = (\omega^2 + \rho^\tau \rho + \rho'^\tau \rho') + (\omega\rho + \omega\rho^\tau) + (\omega\rho' + \omega\rho'^\tau) + (\rho^\tau \rho' + \rho'^\tau \rho)$$

where the terms have been collected to give the components of $\psi^\tau\psi$ according to the direct decomposition $M = K \oplus K\rho \oplus K\rho' \oplus K\rho\rho'$. Now if r is to commute with w_ψ, then $\psi^\tau\psi$ must be in Z. Hence the last three components in (4.4) vanish. As $\omega \neq 0$, $\rho^\tau = -\rho$ and $\rho'^\tau = -\rho'$. But then $(\rho^\tau\rho' + \rho'^\tau\rho) \neq 0$. Thus $\psi^\tau\psi$ is not in Z and r cannot commute with w_ψ.

If r anticommutes with w_ψ, we have shown that $\psi^J = -\xi\psi^\tau\xi^{-1}\epsilon$. Since ψ is in $Z(\rho, \rho')$, $\xi\psi\xi^{-1} = \psi$. Also $\rho^J = \rho^{-1} = \rho\gamma^{-1}$ and $\rho'^J = \rho'^{-1} = \rho'\beta^{-1}$. So we now have

$$\rho\gamma^{-1} + \rho'\beta^{-1} + \omega^J = -(\rho^\tau\epsilon + \rho'^\tau\epsilon + \omega^\tau\epsilon). \tag{4.5}$$

The antiautomorphism J when restricted to Z is either the identity automorphism or an automorphism of order 2. Let Z_0 be the fixed subfield of Z by J; then either $Z = Z_0$ or $Z = Z_0 + Z_0\chi$ where $\chi^J = -\chi$. In the first case, $\omega^J = \omega$. Furthermore $\omega^\tau = \omega$ as ω is in K; so from (4.5), $\epsilon = -1$. In the second case, $\omega = \omega_1 + \omega_2\chi$ where ω_1 and ω_2 are in Z_0. Now (4.5) yields $\epsilon = \pm 1$. Also from (4.5), we have that $\rho\gamma^{-1} = -\rho^\tau\epsilon$ and $\rho'\beta^{-1} = -\rho'^\tau\epsilon$. So if $\tau = 1$ or $\tau = \sigma\sigma'$, then $\beta = \gamma = \pm\epsilon$; this is impossible as γ is a square in L while β is not. If $\tau = \sigma$ or $\tau = \sigma'$, then $\beta = -\gamma = \pm\epsilon$. But $\epsilon = \pm 1$; so one of β or γ equals 1, which is a square in K, contrary to the choice of u and v. This means that r cannot anticommute with w_ψ, which is in $c(u, v)$. As we have also shown that r cannot commute with w_ψ, the assumption that r anticommutes with w yields a contradiction.

This means that r commutes with every 1-involution in $U_q(M, h)$ and hence leaves every 1-dimensional nonisotropic subspace of G invariant. Then by Lemma 1.5, $r(x) = x\delta$ with δ in M for all x in G. If r is linear, then δ is in $Z(\rho, \rho')$. If r is semilinear with respect to an automorphism $\pi = \sigma$, σ', or $\sigma\sigma'$ of M, then π, to be an inner automorphism, must leave $Z(\rho, \rho')$ elementwise invariant, which it does not. Hence δ is in $Z(\rho, \rho')$ and $\delta^2 = \epsilon$ in Z. Again applying the result of Section 1.10 twice, we obtain that $\delta = \delta_0$, $\delta = \delta_1\rho$, $\delta = \delta_2\rho'$, or $\delta = \delta_3\rho\rho'$ with δ_i in Z. As r is a homothetic mapping in $U_q(M, h)$, it follows that $\delta^J\delta = 1$ and hence that $\delta_i^J\delta_i = 1$. Then one of r, ur, vr, or uvr is a homothetic transformation h_η in $U_q(M, h)$, where η is in C. So h_η can be considered in C_n and r is in one of the cosets 1^*, u^*, v^*, or u^*v^*. Therefore, we have the following result.

PROPOSITION 4.4. *Let v be a projective involution of the fourth kind in $U_m(L, g)$. If $n > 4$, $\nu(u^*, v^*) = 4$.*

Now we treat the case that v considered as a projective involution of

$U_n(K,f)$ anticommutes with u. Then v considered as an element of $\Gamma U_m(L,g)$ is a projective semilinear involution relative to the automorphism σ of L. If $v^2=\beta$ where β is in C and if there exists a vector x in F such that $v(x)=x\mu$, then $v^2(x)=x\mu\mu^\sigma$. We therefore distinguish the two cases: β is of the form $\mu\mu^\sigma$ with μ in L (Case I) and β is not of the form $\mu\mu^\sigma$ with μ in L (Case II).

4. 5. Projective semilinear involutions in $\Gamma U_m(L,g)$. Case I. Let v be a projective semilinear involution in $\Gamma U_m(L,g)$ relative to the automorphism σ of L such that $v^2=\beta$ and $\beta=\mu\mu^\sigma$ with μ in L. Then as in Section 1. 8, form the subsfield L_1 of L consisting of the elements of L left fixed by the automorphism $\sigma'=(\mu^{-1})^*\sigma$. $L=L_1\oplus L_1\theta$, where θ is an arbitrary element of L such that $\theta^{\sigma'}=-\theta$. In particular, we pick $\theta=\rho$. By the argument in Section 1. 8, it follows that the center Z_1 of L_1 is contained in $Z(\rho)\cap L_1$. But also $Z\subseteq Z_1$, and as ρ is not in L_1, it follows that $Z_1=Z$.

Now F when considered as a L_1-vector space decomposes into the direct sum $V^+\oplus V^-$ of the eigenspaces of v. Set $F_1=V^+$ and form the group $\Gamma U_m(L_1,g_1)$ acting on F_1 as in Section 1. 8. An element t which commutes with both u and v can be restricted to be in $U_m(L_1,g_1)$ or $Sp_m(L_1,g_1)$, and conversely. If t anticommutes with u and commutes with v, then vt commutes with both u and v and so can be restricted to F_1. Similarly in the other cases ut or uvt can be restricted to F_1.

In case g_1 is herimitian, let w_1 be a 1-involution in $U_m(L_1,g_1)$. Then the extension w of w_1 to E is then a 2-involution in $c(u,v)$. Then if r is in $c(c(u,v))$ and r' is that one of r, ur, vr, or uvr which commutes with u and v, r and r' commute with w. The restriction r_1 of r' to F_1 commutes with w_1. Thus it follows that r_1 leaves invariant every 1-dimensional subspace of F_1.

In case g_1 is alternating, let x be a vector of F_1. Then x lies in a maximal totally isotropic subspace v; and, corresssponding to V, there exists a complementary totally isotropic subspace V' (cf. Dieudonné [4; p. 6]). Now it is clear that we can apply the argument in Section 4. 2 to conclude that $r_1(x)=x\delta_x$.

Consequently, by Lemma 1. 5 or Lemma 1. 6, we can conclude that $r_1(x)=x\delta$ with δ in Z for all x in F_1. Furthermore, as F_1 is nonisotropic relative to g_1, $\delta^T\delta=1$, where T is the antiautomorphism of L_1 corresponding to g_1. Since δ is in Z, $\delta^T=\delta^J$; so δ is in C and the extension r' of r_1 to E is in C_n. Thus r is in one of 1^*, u^*, v^*, or u^*v^*.

PROPOSITION 4. 5. *Let v be a projective semilinear involution in*

$\Gamma U_m(L, g)$ relative to the automorphism σ of L such that $v^2 = \beta$ and β is of the form $\mu\mu^\sigma$ with μ in L. If $n > 4$, then $\nu(u^*, v^*) = 4$.

4.6. Projective semilinear involutions in $\Gamma U_m(L, g)$. Case II. Let v be a projective semilinear involution of $\Gamma U_m(L, g)$ relative to the automorphism σ of L such that $v^2 = \beta$, where β is not of the form $\beta = \mu\mu^\sigma$ with μ in L. Then, as in Section 1.9, adjoin a square root ρ' of β to L forming the sfield $M = L(\rho') = K(\rho, \rho')$. If ϕ is in L, the relation $\phi\rho' = \rho'\phi^\sigma$ holds; in particular $\rho\rho' = -\rho'\rho$. Let Z_M be the center of M. Then clearly $Z \subseteq Z_M \subseteq Z(\rho)$ as $Z(\rho)$ is the center of L. Furthermore, ρ is not in Z_M; and since $[Z(\rho) : Z] = 2$, $Z_M = Z$.

Now form the M-space G by setting $x\rho' = v(x)$ and extend the anti-automorphism J to M by setting $\rho'^J = \rho'^{-1}$. Then, as in Section 1.9, extend the form g to a hermitian sesquilinear form h defined on G and form the group $\Gamma U_q(M, h)$, where $q = m/2 = n/4$. An element t of $U_n(K, f)$ which commutes with both u and v can be considered to be in $U_q(M, h)$ and conversely. In the other cases, ut, vt, or uvt will commute with both u and v; and so one of them can be considered to be in $U_q(M, h)$.

A 1-involution w_1 in $U_q(M, h)$ can then be extended to a 4-involution w in $U_n(K, f)$ which commutes with u and v. Then r in $c(c(u, v))$ commutes or anticommutes with w. If $n > 8$, it follows immediately that r must commute with w. If $n = 8$, the argument of Section 4.4 shows that r must commute with w.[8] Let r' be that one of r, ur, vr, or uvr which commutes with both u and v. Then r' commutes with w and its restriction r_1 to G is a linear projective involution commuting with w_1. That is, r_1 commutes with every 1-involution of $U_q(M, h)$ and so must leave invariant every 1-dimensional nonisotropic subspace of G. Then by Lemma 1.5, $r_1(x) = x\delta$ with δ in Z for every vector x of G. Since r_1 is in $U_q(M, h)$, $\delta^J\delta = 1$ and the extension r' of r_1 to E is a homothetic mapping of E. Thus r' is in 1^*, and r is in one of 1^*, u^*, v^*, or u^*v^*.

---

[8] The argument of Section 4.4 should be altered at two points. First, it should be remarked that τ can be taken to be the identity automorphism, for we can replace r by that one of r, ur, vr, or uvr which commutes with u and v. The extension of this transformation to $U_q(M, h)$ is linear. Second, we should remark that it is possible to choose $\xi \neq \pm 1$ by replacing, if necessary, ξ by $\alpha^J\xi\alpha = h(x\alpha, x\alpha)$ where $\alpha = \rho + \rho'$ or $\alpha = 1 + \rho + \rho'$. Then take $\phi = \xi$ so that $\psi = \xi$ or $\xi \pm 1$. Hence $\psi \neq 0$ and $\psi^J = \psi$. If $\psi^2 = \epsilon^{-1}$, it follows that $\psi^2 = \pm 1$ as $\epsilon^J\epsilon = 1$. But we can assume that $\psi^2 \neq 1$ by a replacement of ξ by $\alpha^J\xi\alpha$ if necessary. So $\epsilon = \epsilon^{-1} = \psi^2 = -1$. If $\psi^2 \neq \epsilon^{-1}$, then $\psi^J = \psi = -\xi\psi\xi^{-1}\epsilon = -\psi\epsilon$. So again $\epsilon = -1$. Now choose $\phi = \rho$. Clearly $\psi^2 \neq -1$; so $\psi^J = \xi\psi\xi^{-1}$. This means that $\rho^J = \xi\rho\xi^{-1}$. Similarly $\rho^J = \xi\rho'\xi^{-1}$ and $(\rho\rho')^J = \xi\rho\rho'\xi^{-1}$. This implies that $(\rho\rho^{\iota})^J = \rho^J\rho'^J = (\rho'\rho)^J$ and thus that $\rho\rho' = \rho'\rho$, which is a contradiction.

PROPOSITION 4.6. *Let v be a projective semilinear involution in $\Gamma U_m(L,g)$ relative to the automorphism σ of L such that $v^2=\beta$, where β is not of the form $\beta=\mu\mu^\sigma$ with μ in L. If $n>4$, then $\nu(u^*,v^*)=4$.*

Propositions 4.1 through 4.6 yield the following Theorem.

THEOREM 4.1. *If the dimension n of E is greater than 4 and if u^* is an involution of the fourth kind in $PU_n(K,f)$, then $\nu_{u^*}=4$.*

V. Characterization of Extremal Involutions.

We now characterize the extremal involutions of $PU_n(K,f)$ when $n>4$ principally by the group-theoretical invariant ω_{u^*}. When $n=6$ and E is commensurable, we use other methods.

5.1. Extremal involutions.

THEOREM 5.1. *If the dimension n of E is greater than 4 and E is not commensurable when $n=6$, then $\omega_{u^*}=8$ for extremal involutions u^* of $PU_n(K,f)$.*

Let u be the 1-involution in u^*. Let v^* and w^* be the distinct involutions of $\mathfrak{I}^*$ commuting with u^* such that $v^*w^*=w^*v^*\neq u^*$, $v^*\neq u^*$, and $w^*\neq u^*$. Let v and w be in v^* and w^*, respectively. By Lemma 2.1, it follows that every projective involution in $c(u,v,w)$ and $c(c(u,v,w))$ is of the first or third kind and commutes with u. By Theorems 2.1 and 2.2, we have that v^* or w^* is an extremal involution if it is of the first kind.

Any element t in $U_n(K,f)$ which commutes with u can be restricted to a transformation t_0 of U^-, the $(n-1)$-dimensional eigenspace of u. As U^- is nonisotropic, the form f can be restricted to a nondegenerate hermitian sesquilinear form f_0 on U^-. The group of restrictions to U^- of elements of $U_n(K,f)$ is then the group $U_{n-1}(K,f_0)$ acting on U^-. The sets $c(u,v,w)$ and $c(c(u,v,w))$ reduce to the sets $c(v_0,w_0)$ and $c(c(v_0,w_0))$ of projective involutions of $U_{n-1}(K,f_0)$. As elements of $U_{n-1}(K,f_0)$, v_0 and w_0 are the same kind of involutions that v and w are. When $n-1>4$, it follows from Theorems 2.1 and 3.1 that $c(c(v_0,w_0))$ contains representatives from only the four cosets 1^*, $v_0{}^*$, $w_0{}^*$, and $v_0{}^*w_0{}^*$ of $PU_{n-1}(K,f_0)$.

A closer investigation shows that this is even true when $n-1=4$. Indeed, this follows from Theorem 2.1 if $v_0{}^*$ or $w_0{}^*$ is an extremal involution. If $v_0{}^*$ and v^* are of the third kind, then $v^2=\beta$ and $v_0{}^2=\beta$, where $\beta=\mu^2$ and μ is not in Z. By Section 1.8, E, considered as a right vector space over K_1, the centralizer of μ in K, is the direct sum of the K-subspaces V^+ and V^-, where $v(x) = \pm x\mu$ for x in V^+ or V^- respectively. Similarly V^-, considered as a K_1-space, is the direct sum $V_0^+ \oplus V_0^-$, where $v_0(x) = \pm x\mu$ for x in V_0^+ or V_0^-, respectively. As in Section 1.8, we form the group $U_n(K_1, f_1)$ acting on the subspace $E_1 = V^+$ and $U_{n-1}(K_1, f_1')$ acting on V_0^+. The form f_1' is the restriction of f_1 to $E_1' = V_0^+$ since

$$E_1 = U^+ \cap V^+ \oplus U^- \cap V^+ = U^+ \cap V^+ \oplus E_1'.$$

If w commutes with v, then by Lemma 2.1, w restricted to E_1 must commute with the restriction u_1 of u to E_1 as u_1 is a 1-involution in $U_n(K_1, f_1)$. Furthermore, then w is of the first or third kinds in $U_n(K_1, f_1)$. So it follows that w_0 is of the first or third kinds in $U_{n-1}(K_1, f_1')$. Hence by Propositions 3.1 and 3.3, $c(c(v_0, w_0))$ again contains projective involutions from only four cosets.

Finally if w anticommutes with v, it is represented by a projective semilinear involution w' in $\Gamma U_n(K_1, f_1)$ relative to an automorphism θ^* of K_1, where θ is an element of K such that $\theta^{\mu^*} = -\theta$. Then as w' commutes with u_1, $w'(U^+ \cap E_1) = U^+ \cap E_1$; so $w(x) = x\lambda$ for some x in $U^+ \cap E_1$. This means that $w'^2 = \gamma$, where $\gamma = \lambda\lambda^{\theta^*}$.

As in Sections 1.8 and 3.4, E_1, when considered as a right vector space over K_2, the fixed subfield of K_1 by $\theta^*\lambda^*$, is the direct sum of the eigenspaces W^+ and W^- of w'. Since w' commutes with u_1 and $U^+ \cap E_1$ is a K_1-subspace, it follows that $U^+ \cap E_1 = U^+ \cap W^+ \oplus U^+ \cap W^-$. If x is in $U^+ \cap W^+$, certainly $f(x, x) \neq 0$. But then, if f_2 is the sesquilinear form defined on W^+ and derived from f_1 in the manner of Sections 1.8 and 3.4, f_2 is but a scalar multiple of the form f restricted to W^+. So $f_2(x, x) \neq 0$. This means that f_2 is not an alternating form. Hence the form f_2' derived from f_1' as in Sections 1.8 and 3.4 is also not alternating inasmuch as f_2' is the restriction of f_2 to $W^+ \cap U^-$. Then by Proposition 3.5 in this last case the projective involutions of $c(c(v_0, w_0))$ come from only the cosets 1^*, v_0^*, w_0^*, and $v_0^* w_0^*$.

Let r^* be an involution of $c(c(u^*, v^*, w^*))$. Then r in r^* can be chosen so that its restriction to U^- is one of 1_0, v_0, w_0, or $v_0 w_0$. Denote by r' that one of 1, v, w, or vw whose restriction to U^- coincides with that of r. Then $r'^2 = r^2 = \epsilon$ in C_n. Furthermore $\epsilon = \delta^2$ where δ is in Z', the center of K', a subfield of K of index 1 or 2 according as r is of the first or third kind. Let R^+ and R^- be the K'-subspaces of r and R'^+ and R'^- be the K'-subspaces of r'. Then since r and r' coincide on U^-, we have $U^- \cap R^+ = U^- \cap R'^+$ and $U^- \cap R^- = U^- \cap R'^-$. Furthermore

$$U^+ = U^+ \cap R^+ \oplus U^+ \cap R^- = U^+ \cap R'^+ \oplus U^+ \cap R'^-.$$

One of these summands may be zero. Indeed, if r and r' are of the first kind, then $U^+ = U^+ \cap R^+$ or $U^+ = U^+ \cap R^-$ and also $U^+ = U^+ \cap R'^+$ or $U^+ = U^+ \cap R'^-$. Therefore for x in U^+, $r(x) = \pm x\delta$ and $r'(x) = \pm x\delta$. Thus there are but two possibilities for $r(x)$ when x is in $U^+ : r(x) = r'(x)$ or $r(x) = -r'(x)$.

If r and r' are of the third kind, then $\dim_{K'} U^+ \cap R'^{\pm} = 1$. Since r is in $c(c(u, v, w))$ while r' is in $c(u, v, w)$, r and r' must commute or anticommute. As they coincide in U^-, they commute. Therefore, $U^+ \cap R^+$ is equal to $U^+ \cap R'^+$, or $U^+ \cap R'^-$, and $U^+ \cap R^-$ is equal to $U^+ \cap R'^-$ or $U^+ \cap R'^+$, respectively. That is, for x in $U^+ = U^+ \cap R^+ \oplus U^+ \cap R^-$, $r(x) = r'(x)$ or $r(x) = -r'(x)$. Consequently $1_0{}^*$, $v_0{}^*$, $w_0{}^*$, and $v_0{}^* w_0{}^*$ are the restrictions of at most 8 involutions in $c(c(u^*, v^*, w^*))$. Since $\omega(u^*, v^*, w^*) \geqq 8$, we have proven the Theorem 5.1.

We shall denote by $\mathcal{J}^*$ the set of involutions of $\mathfrak{I}^*$ for which $\omega_{u^*} = 8$.

5.2. Involutions of the second kind. We assume that $n > 4$. It can be shown that the involutions of the second kind are not in $\mathcal{J}^*$ providing $n > 6$. To treat the case $n = 6$, we consider the set $\mathcal{K}^*$ of involutions of $PU_n(K, f)$ each of which is not conjugate to the product of itself with two other distinct commuting involutions with which it commutes and to which it is conjugate.

THEOREM 5.2. *Let the dimension of E be greater than 4. Then the set $\mathcal{K}^*$ contains the extremal involutions of $PU_n(K, f)$ but not the involutions of the second kind.*

Let u^*, v^*, w^* be distinct commuting and conjugate extremal involutions. Then by Section 1.6, $u^* v^* w^*$ is a 3-involution and hence not conjugate to an extremal involution.

Let u^* be an involution of the second kind in $PU_n(K, f)$. Pick u to be an involution of $U_n(K, f)$ in u^* with totally isotropic eigenspaces U^+ and U^- of dimension $m = n/2$. Dieudonné [2; p. 7] has shown that if $U_0{}^+$ and $U_0{}^-$ are any two maximal totally isotropic subspaces whose sum $U_0{}^+ \oplus U_0{}^- = E$, then there exists an element t of $U_n(K, f)$ mapping U^+ onto $U_0{}^+$ and U^- onto $U_0{}^-$. The proof of Dieudonné is stated for the symplectic groups but applies in this case as well. Thus if $u^2 = \gamma$, where $\gamma = \lambda^2$, we define u_0 such that $u_0(x) = \pm x\lambda$ according as x is in $U_0{}^+$ or $U_0{}^-$. Then u and u_0 are conjugate in $U_n(K, f)$. Since every involution $u_0{}^*$ of the second kind in $PU_n(K, f)$ contains an involution u_0 as defined above, u^* is conjugate to the involution $u_0{}^*$.

Now pick a basis $x_1, x_2, \cdots, x_m$ for U^+ and $x_{m+1}, x_{m+2}, \cdots, x_n$ for U^- such that $f(x_i, x_j) = 0$ except when $j - i = \pm m$, in which case $f(x_i, x_j) = 1$. Let $P_1 = x_1K$, $P_2 = x_2K$, $P_3 = x_3K \oplus \cdots \oplus x_mK$, $P_4 = x_{m+1}K$, $P_5 = x_{m+2}K$ and $P_6 = x_{m+3}K \oplus \cdots \oplus x_nK$. As $m \geqq 3$, all subspaces P_i are not zero. Let

$$V^+ = P_1 \oplus P_2 \oplus P_6,\; V^- = P_4 \oplus P_5 \oplus P_3,$$
$$W^+ = P_1 \oplus P_5 \oplus P_3,\; W^- = P_2 \oplus P_4 \oplus P_6.$$

Then V^+, V^-, W^+, and W^- are all totally isotropic since the summands P_i in the decomposition of these spaces are all orthogonal. Let v and w be the projective involutions of the second kind in $U_n(K, f)$ whose eigenspaces are $V^\pm$ and $W^\pm$, respectively, and which are conjugate to u. Then u, v, and w all commute and $u^2 = v^2 = w^2 = \beta$. Thus $u_0 = uvw$ is a projective involution of the second kind as $u_0{}^2 = \beta^3$. Hence $u_0{}^*$, v^*, and w^* are conjugate to u^*. Since u^*, v^*, and w^* commute, Theorem 5.2 is proved.

5.3. Involutions of the third kind. Let now u^* be an involution of the third kind in $PU_n(K, f)$ and assume that $n > 4$. We shall show that $\omega_{u^*} > 8$. Let u be a representative of u^*. Then $u^2 = \gamma$ with γ in C and $\gamma = \lambda^2$, where λ is not in Z. As in Section 1.8, let K_1 be the fixed subsfield of K by λ^*. Then $K = K_1 \oplus K_1\theta$, where θ is an element of K which anticommutes with λ. Then E can be considered as a K_1-space E^*, and $E^* = U^+ \oplus U^-$, where U^+ and U^- are the positive and negative eigenspaces of u. Set $E_1 = U^+$ and form the group $\Gamma U_n(K_1, f_1)$ acting on E_1.

If $\mathcal{S}^*$ is a set of involutions in $PU_n(K_1, f_1)$, let $c^*(\mathcal{S}^*)$ denote those involutions of $PU_n(K_1, f_1)$ which commute with every involution of $\mathcal{S}^*$. If $v_1{}^*$ and $w_1{}^*$ are distinct commuting involutions in $PU_n(K_1, f_1)$, denote the order of $c^*(c^*(v_1{}^*, w_1{}^*))$ by $\nu^*(v_1{}^*, w_1{}^*)$.

By Theorem 2.2, Section 2.3, there exist two involutions of the first kind in $PU_n(K_1, f_1)$, which we here denote by $v_1{}^*$ and $w_1{}^*$ such that $\nu(v_1{}^*, w_1{}^*) > 4$. So we can find an involution $r_1{}^*$ of the first kind (which is one of the $r_i{}^*$ of Section 2.3) in $c^*(c^*(v_1{}^*, w_1{}^*))$ which is not one of 1^*, $v_1{}^*$, $w_1{}^*$ or $v_1{}^*w_1{}^*$. The restriction of u to E_1 is a homothetic mapping h_λ in the center $C_n{}^*$ of $U_n(K_1, f_1)$. Let v_1', w_1' and r_1' be the p-involutions ($p \leqq [n/2]$) contained in $v_1{}^*$, $w_1{}^*$, and $r_1{}^*$, respectively; set $v_1 = h_\lambda v_1'$ and $w_1 = h_\lambda w_1'$. We still have v_1 in $v_1{}^*$ and w_1 in $w_1{}^*$; so $r_1{}^*$ is in $c^*(c^*(v_1{}^*, w_1{}^*))$.

We can extend r_1 to an involution r of $U_n(K, f)$ by setting $r(x\theta) = r_1(x)\theta$ for x in $E_1 = U^+$. This defines $r(x)$ on $U^- = U^+\theta$ and consequently on E as $E^* = U^+ \oplus U^-$. Similarly, we can extend v_1 and w_1 to projective involutions

v and w in $U_n(K,f)$. But now $v^2=w^2=\gamma$; so v and w are projective involutions of the third kind in $U_n(K,f)$. Therefore, by Theorem 3.3, v and w are in $\mathfrak{J}$.

Since h_λ, v_1 and w_1 all commute, their extensions u, v, and w must all commute. So form $c(c(u,v,w))$. We shall show that r is in $c(c(u,v,w))$. It will then follow that $\omega(u^*,v^*,w^*)>8$, as the restriction r_1 of r to E_1 does not coincide with the restriction of any of the elements of 1^*, u^*, v^*, w^*, u^*v^*, v^*w^*, w^*u^*, and $u^*v^*w^*$, which are in the cosets 1^*, $v_1{}^*$, $w_1{}^*$ and $v_1{}^*w_1{}^*$ of $PU_n(K_1,f_1)$.

If t is a projective involution in $c(u,v,w)$ which commutes with u, then the restriction t_1 of t to E_1 will be a projective involution in $c(v_1,w_1)$. By Section 2.3, $t_1r_1=r_1t_1$. Let x be in U^+ so that $x\theta$ is in U^-. Then $t(r(x\theta))=t_1(r_1(x))\theta=r_1(t_1(x))\theta=r(t(x\theta))$. Since $E=U^+\oplus U^-$, t commutes with r. If t is a projective involution in $c(u,v,w)$ which anticommutes with u, then set $t'=h_\theta t$. Thus t' will commute with u and commute or anticommute with v and w. The restriction t_1 of t' to E_1 will then commute or anticommute with v_1 and w_1. But the subspaces P_i were chosen in Section 2.3 so that no transformation could anticommute with v_1 or w_1. So t_1 commutes with v_1 and w_1 and hence leaves the eigenspaces of r_1 invariant. Thus $t_1r_1=r_1t_1$. Again let x be in U^+ so that $x\theta$ is in U^-. Then $t'(r(x\theta))=t_1(r_1(x))\theta=r_1(t_1(x))\theta=r(t'(x\theta))$. This means that t' commutes with r and so t commutes with r. Thus r is in $c(c(u,v,w))$ and $\omega(u^*,v^*,w^*)>8$.

THEOREM 5.3. *If the dimension of E is greater than 4 and if u^* is an involution of the third kind in $PU_n(K,f)$, $\omega_{u^*}>8$.*

5.4. Involutions of the fourth kind. We assume that $n>4$ and E is not commensurable when $n=6$. We can show that $\mathcal{J}^*$ contains no involutions of the fourth kind when $n>12$. However, the following argument yields a more general result.

Theorems 5.1, 5.2, and 5.3 imply that $\mathcal{J}^*\cap\mathcal{K}^*$ contains only the extremal involutions and possibly the involutions of the fourth kind. We denote by $\mathfrak{M}(u^*)$ a maximal set of mutually commuting involutions of $\mathcal{J}^*\cap\mathcal{K}^*$ containing u^*. By Lemma 2.1, every involution of $\mathfrak{M}(u^*)$ is extremal or none is. By Section 1.6, every maximal set $\mathfrak{M}(u^*)$ of extremal involutions contains n elements. Furthermore the product of any two or three such involutions is a 2- or 3-involution. Consequently these products are not in $\mathfrak{J}^*$ by virtue of Theorem 2.2. We therefore define $\mathcal{L}^*$ to be the subset of $\mathcal{J}^*\cap\mathcal{K}^*$ consisting of the elements u^* such that the sets $\mathfrak{M}(u^*)$

always possess n elements and such that the product of any two or three involutions in a set $\mathfrak{M}(u^*)$ is not in $\mathfrak{J}^*$. Our principal result is the following theorem.

THEOREM 5.4. *If the dimension n of E is greater than 4 and E is not commensurable when $n = 6$, then the set $\mathcal{L}^*$ is the set of extremal involutions.*

In view of the preceding remarks, it shall suffice to show that if u^* is an involution of the fourth kind and $\mathfrak{M}(u^*)$ contains n elements, then there exists a product of 2 or 3 involutions of $\mathfrak{M}(u^*)$ in $\mathfrak{J}^*$. Let $u^* = u_1^*, u_2^*, \cdots, u_n^*$ be the elements of $\mathfrak{M}(u^*)$ and u_i be representatives for u_i^*, $i = 1, 2, \cdots, n$. Then $u_iu_j = \pm u_ju_i$, and $u_i^2 = \gamma_i$, where γ_i is not a square in K. Set $(u_iu_j)^2 = \gamma_{ij} = \pm \gamma_i\gamma_j$. First consider that there is a pair i and j, $i \neq j$, such that γ_{ij} is not a square in Z. Then u_iu_j is a projective involution of the third or fourth kinds and is consequently in $\mathfrak{J}$ by Theorems 3.1 and 4.1.

On the other hand, suppose $\gamma_{ij} = \lambda_{ij}^2$ with λ_{ij} in Z for all i and j, where $i \neq j$. If then there exists a projective involution u_k, $k > 2$, which commutes with u_1u_2, then $u_1u_2u_k$ is of the fourth kind as $\lambda_{12}^2\gamma_k$ cannot be a square in K. Finally suppose that all u_k, $k > 2$, anticommute with u_1u_2. Since $n > 4$, there exist two u_k, say u_3 and u_4 which commute with one, say u_1, and anticommute with the other. But then u_1 commutes with u_3u_4, and by the previous argument $u_1u_3u_4$ is in $\mathfrak{J}$. This proves the theorem.

5.5. E commensurable and $n = 6$. Designate by π_{u^*} the maximum of the numbers of involutions in the sets $\mathfrak{N}(u^*)$ of mutually commuting and conjugate involutions containing u^*.

THEOREM 5.5. *If E is commensurable and $n = 6$, then $\pi_{u^*} = 6$ for extremal involutions u^* in $PU_n(K, f)$ and $\pi_{u^*} > 6$ for nonextremal involutions u^* of the first kind.*

Because E is commensurable, all extremal involutions are conjugate to each other. By Section 1.6, the maximum number of such involutions in a set of mutually commuting involutions is 6. Therefore $\pi_{u^*} = 6$ if u^* is extremal.

Similarly, all 2-involutions [3-involutions] of $PU_n(K, f)$ are conjugate. Dieudonné [5; p. 5] has shown that there exist 15 2-involutions [20 3-involutions] in a maximal set of commuting 2-involutions [3-involutions] of $GL_n(K)$. The same result is true for $U_n(K, f)$ by the same argument. These 2-involutions [3-involutions] come from 15 [10] distinct cosets as there is but one [two] involutions in a coset which is a 2-involution [3-involution] of $PU_n(K, f)$. Thus $\pi_{u^*} > 6$ when u^* is a 2-involution [3-involution].

THEOREM 5.6. *When the dimension n of E is 6 and E is commensurable, then the set of extremal involutions in $PU_n(K,f)$ can be characterized as the set of involutions u^* for which*

(a) $\pi_{u^*} = 6$.

(b) *The product of u^* with any two other involutions from any set $\mathfrak{N}(u^*)$ of mutually commuting involutions conjugate to u^* is never conjugate to u^*.*

(c) *Any product of k distinct involutions from $\mathfrak{N}(u^*)$ is conjugate to any other product of k distinct involutions from $\mathfrak{N}(u^*)$.*

(d) *The product of all the involutions in $\mathfrak{N}(u^*)$ is 1^*.*

(e) *For any three involutions v_1^*, v_2^*, and v_3^* in $\mathfrak{N}(u^*)$, $\omega(v_1^*, v_2^*, v_3^*) = 8$.*

We first mention that extremal involutions u^* satisfy the above conditions. Indeed, Theorem 5.5 guarantees (a) is satisfied. Theorem 5.2 justifies (b). From Section 1.6, we see that any product of k distinct extremal involutions is a p- or $(n-p)$-involution, and from Section 1.5 all such involutions are conjugate. The justification for (d) is in Section 1.6. Finally, in the proof of Theorem 5.1, we established (e); the assumption that E be not commensurable was used there only to exclude 3-involutions from the set $\mathcal{J}^*$ from which we chose the involutions v^* and w^* commuting with u^*.

Theorem 5.5 states that the 2- and 3-involutions of $PU_6(K,f)$ do not satisfy (a). Theorem 5.2 states that the involutions of the second kind in $PU_6(K,f)$ do not satisfy (b). So we must only consider the cases where u^* is an involution of the third or fourth kind. We show that not all the conditions of Theorem 5.6 are satisfied.

We assume that u^* satisfies (a), (b), (c), and (d). Thus since $\pi_{u^*} = 6$, we can pick a set $\mathfrak{N}(u^*)$ of 6 mutually commuting involutions conjugate to u^*, one of which is u^*. Let $u_1^* = u^*, u_2^*, \cdots, u_6^*$ be the elements of $\mathfrak{N}(u^*)$. Pick representatives u_i in u_i^*. By (d), we have that $u_1u_2u_3u_4u_5u_6 = h_\lambda$ in C_n. This means that u_1 commutes with $u_2u_3u_4u_5u_6$ and thus commutes with one u_i, $i > 1$, say u_2. It then follows that $w_1 = u_1u_2$ is a projective involution of the first kind. Consequently by (c), $w_2 = u_3u_4$ and $w_3 = u_5u_6$ are in cosets w_2^* and w_3^*, respectively, which are conjugate to w_1^* when considered as elements of $PU_6(K,f)$. Since $w_1^*w_2^*w_3^* = 1^*$, it is possible to choose the representatives w_i and u_i so that w_i, $i = 1, 2, 3$, are involutions of $U_n(K,f)$ and $w_1w_2w_3 = 1$.

We shall show that the involutions w_i are 2-involutions. Indeed, as they come from conjugate cosets, they are all either 1-involutions, 2-involutions or 3-involutions. However, we have that $w_1 = w_2 w_3$.

In the case that the involutions w_i are 1-involutions or 3-involutions, $w_2 w_3$ is easily verified to be a 2-involution while $w_1 = w_2 w_3$. So it follows that w_i are all 2-involutions. Then by (c) $u_i^* u_j^*$ is a 2-involution for any two distinct involutions u_i^*, u_j^*. Thus $u_i u_j$ is a projective involution of the first kind in $U_n(K,f)$ which is in the coset of a 2-involution; that is, $u_i u_j$ have a 2-dimension eigenspace. As $n = 6$, no element t in $U_6(K,f)$ anti-commutes with $u_i u_j$. This means that all the projective involutions u_i commute; for, should $u_i u_j = -u_j u_i$, then would $u_i(u_i u_j) = -(u_i u_j) u_i$.

Let now u^* be an involution of the third kind, and form the group $\Gamma U_6(K_1, f_1)$ acting on the K_1-subspace of E defined in Section 1.8. Since $u_i u_1$, $i > 1$, is in the coset of a 2-involution of $PU_6(K,f)$, u_i is represented in $U_6(K_1, f_1)$ by a projective involution u_i' of the first kind with a 2-dimensional eigenspace U_i^+ of E_1 and 4-dimensional eigenspace U_i^-. As there are five such involutions u_i', we can find two, say u_2 and u_3, such that $U_2^+ \cap U_3^+$ is 1-dimensional. It then follows that none of the spaces $U_2^\pm \cap U_3^\pm$ are zero. Hence by the argument of Section 5.3, $\omega(u_1^*, u_2^*, u_3) > 8$. Thus u^* does not fulfill condition (e).

Now assume that u^* is of the fourth kind in $PU_6(K,f)$. As in Section 1.9, form the group $\Gamma U_3(L,g)$ acting on the L-vector space F. In $U_6(K,f)$ the projective involutions $u_i u_1$ have distinct 2- and 4-dimensional eigenspaces. Considered as elements u_i' of $U_3(L,g)$, they are 1-involutions as the L-dimension of a subspace is half its K-dimension. Furthermore, u_i' and u_j' always represent distinct cosets u_i^* and u_j^* from $PU_3(L,g)$; for, otherwise, their eigenspaces would coincide as subspaces of F. This would mean that the eigenspaces of $u_i u_1$ and $u_j u_1$ would coincide as subspaces of E. Then we would have $u_i^* u_1^* = u_j^* u_1^*$; that is, $u_i^* u_j^* = 1^*$, which is contradiction to (c).

This implies that $u'_2{}^*, u'_3{}^*, \cdots, u'_6{}^*$ form a set of distinct mutually commuting extremal involutions in $PU_3(L,g)$. Since the maximum number of such involutions is 3 by Section 1.6, we have a contradiction. This establishes Theorem 5.6.

UNIVERSITY OF MICHIGAN AND UNIVERSITY OF WASHINGTON.

REFERENCES.

[1] G. Birkhoff and J. von Neumann, "The logic of quantum mechanics," *Annals of Mathematics*, vol. 37 (1936), pp. 823-243.

[2] J. Dieudonné, "La géometrie des groupes classiques," *Ergebnisse der Mathematik und ihrer Grenzgebeite*, Neue Folge, Heft 5 (1955), Berlin (Springer-Verlag).

[3] ———, "Les extension quadratiques des corps non commutatifs et leurs applications," *Acta Mathematica*, vol. 87 (1952), pp. 175-242.

[4] ———, "Sur les groupes classiques," *Actualités Scientifiques et Industrielles*, No. 1040 (1950), Paris (Hermann et Cie).

[5] ———, "On the automorphisms of the classical groups," *Memoirs of the American Mathematical Society*, No. 2 (1950).

[6] G. W. Mackey, "Isomorphisms of normed linear spaces," *Annals of Mathematics*, vol. 43 (1942), pp. 244-260.

[7] C. E. Rickart, "Isomorphic groups of linear transformations," *American Journal of Mathematics*, vol. 72 (1950), pp. 451-464.

[8] ———, "Isomorphic groups of linear transformations II," *American Journal of Mathematics*, vol. 73 (1951), pp. 697-716.

[9] ———, "Isomorphisms of infinite-dimensional analogues of the classical groups," *Bulletin of the American Mathematical Society*, vol. 57, No. 6 (1951), pp. 435-448.

ON THE LOCAL EMBEDDING PROBLEMS IN THE DIFFERENTIAL GEOMETRY OF SURFACES.*

By AUREL WINTNER.

Let D be an open, simply connected domain in a (u^1, u^2)-plane. For the purposes at hand, it can always be assumed that D is sufficiently small, in the sense of being a sufficiently small neighborhood of a given point, say of $(u^1, u^2) = (0, 0)$. If the elements $g_{ik} = g_{ik}(u^1, u^2)$ of a binary, symmetric matrix (g_{ik}) of non-vanishing determinant are given on D as functions of class C^2, then

$$g_{ik}(u^1, u^2)\,du^i du^k \tag{1}$$

will be called a C^2-*metric* (on D). Then, since $g^{ik} = (g_{ik})^{-1}$ and the first and second derivatives of $g_{ik}(u^1, u^2)$ exist and are continuous, (1) has a continuous curvature $K = K(u^1, u^2)$. Suppose that, for some reason, the curvature K turns out to be of class C^1 (rather than just continuous) on D. Then the C^2-metric (1) will be called a *regular* C^2-*metric* (on D).

It was shown in [12] that, whereas certain theorems in differential geometry are false for C^2-metrics, they become true for regular C^2-metrics. The purpose of this note is to show that the notion of a regular C^2-metric (or of the corresponding notion of a regular C^n-metric, where $n \geqq 1$) leads to improved formulations of certain solved, and to corresponding sharp formulations of certain unsolved, problems in the theory of local embeddings.

It will be assumed throughout that the metric (1), instead of being just of non-vanishing determinant, is *positive definite*, that is,

$$\det g_{ik} > 0 \text{ and } g_{ii} > 0. \tag{2}$$

Let S be a surface of class C^3 in an (x, y, z)-space and let S be a "small," simply connected, open piece of a surface ("small" in the same sense in which the domain D was defined to be "small"). Then S has a one-to-one parametrization,

$$S\colon X = X(u^1, u^2), \qquad X = (x, y, z), \tag{3}$$

which is a C^3-parametrization (so that (u^1, u^2) varies over a D on which

---

* Received July 27, 1955.

the vector function X is of class C^3 and has, with respect to u^1 and u^2, a (3 by 2) Jacobian matrix of rank 2 or, what is the same thing, the first fundamental form

$$|\, dX(u^1, u^2)\,|^2 \equiv g_{ik}(u^1, u^2)\, du^i du^k \tag{4}$$

is of non-vanishing determinant on D). Only such parametrizations (3) of the C^3-surface will be used as satisfy these conditions. Corresponding definitions and conventions will be used in connection with C^n-surfaces S for any n.

The considerations will center about the following fact:

(i) *If* (3) *is a C^3-surface, then its first fundamental form* (4), *instead of being just a C^2-metric, is a regular C^2-metric.*

The truth of this fact (i) becomes clear if the *theorema egregium* is read in the direction opposite to that emphasized by Gauss (cf. [12], p. 686).

If a metric (1), rather than a surface S, is given, then by an *embedding of* (1) is meant any surface (3) satisfying (4) (provided that there exists such a surface). Clearly, (i) can be interpreted as follows: In order that a metric admits some C^3-embedding, it is necessary that the metric be a regular C^2-metric. It is natural to ask whether this necessary condition is sufficient as well, at least if the given metric has no parabolic points, that is, if the curvature $K(u^1, u^2)$ of the given regular C^2-metric is either positive throughout or negative throughout. It turns out that this converse of (i) is false, even if

$$K(u^1, u^2) \neq 0 \text{ on } D. \tag{5}$$

The failure of the converse of (i) would be less surprising if it were not claimed even under the assumption (5). In fact, if $K(u^1, u^2)$ changes sign on D, then the Monge-Ampère equation, on which the problem of embedding depends, is in a state of transition from the elliptic ($K > 0$) to the hyperbolic ($K < 0$) type of a partial differential equation of second order, and there are no general existence theorems in such a situation of the Tricomi type (except in the analytic case of Cauchy-Kowalewski). A corresponding remark holds if the metric (1), given on D, fails to satisfy (5) but is either non-hyperbolic ($K \geqq 0$) or non-elliptic ($K \leqq 0$), without being parabolic ($K \equiv 0$).

What *is* true in the case (5) is the following *partial* converse of (i):

(ii) *Every regular C^2-metric* (1) *which is of non-vanishing curvature (that is, which is either hyperbolic,*

$$K(u^1, u^2) < 0, \tag{6}$$

or elliptic,

$$K(u^1, u^2) > 0, \tag{7}$$

on D) *possesses* C^2*-embeddings* S (*locally*; that is, embeddings of D_0, if D_0 is a sufficiently small neighborhood of any given point of D).

What *is not* true is a certain *full* converse (i) in the case (5), since the situation as follows:

(iii) *Under the assumptions of* (ii) *on the metric* (1), *a* C^2*-embedding* S, *the existence of which is assured by* (ii), *can fail to be a* C^3*-embedding not only in the hyperbolic case but also in the elliptic case.*

It will remain *undecided* whether a *complete* converse of (i) in the case (5) is false in the following sense also:

(iii ?) *Under the assumptions of* (ii) *on the metric* (1), *must there exist some* C^3*-embedding* S *in the hyperbolic case and/or in the elliptic case?*

Actually, the answer to the questions (iii ?) does not seem to be known even if (1) is not restricted by the assumption (5). But if (5) is omitted, then nothing is known, except in the analytic case. In fact, even the following question (representing the weakest form of anything like (ii) when (5) is not assumed) is unanswered:

(ii ?) *Does every* C^∞*-metric possess some* C^∞*-embedding or, at least, some* C^2*-embedding?*

It is understood that (1) is called a C^∞-metric if the three functions $g_{ik}(u^1, u^2)$ possess continuous partial derivatives of arbitrarily high order (then the same is true of $K(u^1, u^2)$; so that every "C^∞-metric" is a "regular C^∞-metric"). The trouble is that there is not available an existence theorem (not even with suitable, rather than with arbitrary, Cauchy data) for a system of partial differential equations which has the same form as a Cauchy-Kowalewski system but in which the coefficient functions, instead of being analytic, are just of class C^∞. As a matter of fact, such an existence theorem seems to be doubtful even in case of *linear* Cauchy-Kowalewski systems with coefficient functions of class C^∞.

In the analytic case, the answer to the analogue of question (ii ?) is known. In fact, every analytic metric (1) possesses analytic embeddings. This follows from the Cauchy-Kowalewski theorem, since the latter can be applied to the problem of embedding so as to avoid any reference to assumption (5).

In the *elliptic* case of question (iii ?), there is an affirmative answer under a "restriction of comparison." This is the content of the following criterion (iv):

(iv) *If an elliptic, regular C^2-metric possesses a C^4-embedding, then all of its C^2-embeddings (which, by* (ii), *exist) are C^3-embeddings.*

It is clear from the case $K \equiv -1$ of an analytic metric that (iv) becomes false if its assumption (7) is replaced by (6). The following question will remain unanswered:

(iv bis ?) *Does* (iv) *remain true if its C^4 is reduced to C^3 ?*

In (iv ?) below, the regularity of a C^1-metric is meant in the sense of the existence of a continuous curvature $K = K(u^1, v^2)$ for the C^1-metric (1), when K is defined in the sense of Weyl [9], pp. 42-44; cf. the comments preceding (8) below.

(iv) suggests a question concerning an elliptic analogue of (ii). But this question remains undecided:

(iv ?) *Must every elliptic, regular C^1-metric possess an embedding which is a surface S of class C^2 ?*

A positive answer to this question (iv ?) would be important, since if the answer were in the negative, it would follow that, *for local reasons,* Weyl's embedding problem *in the large* cannot be solved with that generality with which he has stated it originally (cf. [6], p. 485 and p. 487, and the footnote in [12], p. 685).

Corresponding to the two cases, (6) and (7), of (ii), there is a hyperbolic counterpart of (iv bis ?). Since no potential-theoretical re-enforcements can be forthcoming in this hyperbolic case, it appears that the answer to this counterpart of (iv bis ?) is in the negative (a counterexample, if any, seems to be harder in the elliptic case).

The following pair of problems is of quite another character.

(+ ?), (— ?) *Let S be a C^1-embedding of a metric* (1) *on D which is an analytic metric or, for that matter, just a regular C^1-metric (so that* (1) *defines a continuous curvature). Must then S be a convex surface or a saddle surface over D according as the curvature of* (1) *satisfies* (7) *or* (6)?

It is understood that convexity and saddle character of C^1-surfaces are defined in terms of the tangent plane (which, locally, supports S in the first case but not in the second). If S is a C^2-surface, then it has a (continuous) second fundamental form, and so the questions are trivial. But if S is of class C^1 without being of class C^2, then (7) or (6) cannot be expressed in terms of a second functional form, since the latter fails to exist.

It will be clear from the proofs that both (ii) and (iii) remain true if C^2 and C^3 are replaced by C^n and C^{n+1}, respectively, where $n \geqq 2$. A corresponding remark holds, of course, in (i), where $n \geqq 1$, whereas the truth of the analogues of (ii) and (iii) will remain undecided if $n = 1$. It is understood that by a regular C^n-metric is meant a C^n-metric possessing a curvature $K(u^1, u^2)$ which is a function of class C^{n-1}.

The general concept and the name of a regular C^1-metric are due to H. Weyl and H. Cartan, respectively; for references, cf. [2]. For metrics (1) which are given in an isothermic form (that is, in a form in which

$$(8) \qquad g_{11} = g_{22} \text{ and } g_{12} = 0$$

hold), the concept of a regular C^1-metric is contained already in a paper of Lichtenstein [8], pp. 87-88 and 96-100 (in fact, cf. the relation to which formulæ (15) in [12] reduces in the isothermic case). This is the more interesting since, as it turned out in the meantime, there is no loss of generality in assuming (8). In fact, it was shown in [2] that, after a suitable (u^1, u^2)-transformation (which is of class C^2 and of non-vanishing Jacobian), *every* regular C^1-metric (1) becomes a regular C^1-metric satisfying (8).

Proof of (ii). Consider Darboux's differential equation (cf. [1], pp. 240-241) for an unknown function $z = z(u, v)$ of class C^2 of the embedding problem, where, if (1) is the given metric (to be embedded),

$$(9) \qquad u = u^1, \qquad v = u^2.$$

Let $p, \cdots, t$ denote the partial derivatives $z_u, \cdots, z_{vv}$. Darboux's differential equation for $z = z(u, v)$ is a Monge-Ampère equation,

$$(10) \qquad g^2 \cdot (rt - s^2) + ar + 2bs + ct + d = 0,$$

in which the coefficients $g^2 = \det g_{ik} > 0$, a, b, c, d do not contain z, are polynomials in p, q, contain the independent variables u, v only in terms of the (given) functions

$$(11) \qquad g_{ik}(u, v);\ g_{ik\,u}(u, v),\ g_{ik\,v}(u, v),\ K(u, v),$$

and are polynomials in the latter. It is also well-known that the bicharacteristic type of (10) is hyperbolic or elliptic according as the metric (1) is hyperbolic or elliptic. In view of (5), this means that the partial differential equation (10) is of hyperbolic or of elliptic type according as

$$(12-) \qquad K(u, v) < 0, \qquad\qquad (12+) \qquad K(u, v) > 0.$$

Since (ii) assumes that (1) is a regular C^2-metric, all functions (11) are of class C^1 in (u, v). Hence, it is possible to apply to (10) Theorem (V), p. 849, of [4] on hyperbolic Monge-Ampère equations in the case (12—) and Theorem (§), pp. 316-317, of [5] on general (not necessarily Monge-Ampère) elliptic equations in the case (12+) (in which case, instead of the C^1-character of the functions (11), their C^λ-character, where λ is a Hölder index satisfying $0 < \lambda < 1$, would suffice; cf. (§), *loc. cit.*). Thus, the pair of existence theorems, just quoted, leads to the following result:

If (1) is a regular C^2-metric on a (u, v)-domain D, where $(u, v) = (u^1, u^2)$, then, corresponding to every non-parabolic point (u_0, v_0) of D, there exists a neighborhood D_0 of (u_0, v_0) on which Darboux's equation (10) posesses solutions $z = z(u, v)$ of class C^2. In addition, since such a C^2-solution $z = z(u, v)$ belongs to every sufficiently smooth choice of Cauchy initial data in the hyperbolic case and of Dirichlet boundary data in the elliptic case (cf. *loc. cit.*), these initial or boundary data can be chosen so "small" that, if the corresponding C^2-solution $z(u, v)$ is subjected to the first Beltrami operator ∇ of the given metric (1), then the resulting function $\nabla z(u, v)$ (which clearly is a function of class C^1 on D_0 and, since (1) is positive definite, attains non-negative values only) will satisfy the inequality

$$\nabla z(u, v) < 1 \tag{13}$$

at every point (u, v) of D_0 (if the neighborhood D_0 of (u_0, v_0) is small enough).

The existence of a surface S having the properties claimed by (ii) now follows, by Darboux's procedure (cf. the first italicized statement in [1], p. 241), from the eixstence of functions $z = z(u, v)$ of class C^2 satisfying both (10) and (13) on D_0.

It should be noted that, in view of that device in Darboux's procedure which is expressed by the italicized statement in [1], p. 240, recourse must be had to a certain lemma on "euclidean" mappings. The lemma in question is, however, contained in the statements (I*), (II*) of [12], p. 681.

Remark. Owing to the freedom in choosing the (smooth) Dirichlet data in the case $K > 0$ above, the C^2-solutions $z = z(u, v)$ of (10) can be chosen so as to satisfy, besides (13), other inequalities also (on D_0, if D_0 is small enough). Since such an inequality is $H^2 \neq K$, where $H = H(u, v)$ denotes the mean curvature of a surface S, to be supplied by (ii), and $K = K(u, v)$ is the Gaussian curvature, hence a function given along with (1), it follows that, *under the assumptions of* (ii), *the C^2-embedding S can be chosen so as to be free of umbilical points.*

Proof of (iii). Let (1) be given by $(du)^2 + (e^{cu}dv)^2$, where $(u, v) = (u^1, u^2)$ and $c = \text{const}$. Then the curvature $K(u^1, u^2)$ is $-c$, hence $K \equiv -1$ if $c = 1$. Since $K < 0$, the Monge-Ampère equation (10) belonging to this analytic metric (1) is of hyperbolic type, i. e., it has distinct real bicharacteristics, and possesses therefore solutions $z(u, v)$ which are of class C^n without being of class C^{n+1}, where $n \geqq 2$ can be chosen arbitrarily (for details, cf. a corresponding construction in [5], pp. 329-330).

Thus (iii) is clear in the hyperbolic case $(K < 0)$. An example proving (iii) in the elliptic case is less straightforward. Such an example $(K > 0)$ was given in [6], p. 486.

Remark. It would be desirable to decide whether such an example $(K > 0)$ is possible also when it is required to be *free of umbilical points* (in fact, a perusal of the example given in [6], p. 486, shows that the condition $H^2 \neq K$ is violated at the crucial point of the surface obtained there).

Proof of (iv). The following concepts will be needed:

Two metrics, say (1) and

$$E(u, v)\,du^2 + 2F(u, v)\,du\,dv + G(u, v)\,dv^2, \tag{14}$$

are *C^m-isometric* (for a fixed $m > 0$) if there exists a local transformation T,

$$u = u(u^1, u^2), \qquad v = v(u^1, u^2), \tag{15}$$

in which the functions u, v are of class C^m in (u^1, u^2), have a non-vanishing Jacobian, and are such that (1) and (14) become identical by virtue of (15) (or, equivalently, of its inverse, T^{-1}). Thus, if (14) is a regular C^n-metric for a fixed $n > 0$ (that is, if the coefficients of (14) are functions of class C^n and if, in addition, (14) possesses a curvature $K = K(u, v)$ which is a function of class C^{n-1}), then (14) is C^{n+1}-isometric to a regular C^n-metric (1) satisfying (8). This was proved for $n > 1$ in [12] and for the (more intricate) case $n = 1$ in [2]. The necessity of dealing with the concept of C^m-isometry was pointed out in [11]; cf. also [10]. The answer to the question raised in [11] was answered by Hartman [3] in the affirmative. It is readily seen that Hartman's result contains the following fact:

LEMMA. *Every C^1-isometry T of two regular C^n-metrics is a C^{n+1}-isometry* ($n > 0$ is arbitrarily fixed).

If C^1 and $n > 0$ in this Lemma are reduced to C^2 and $n > 1$, respectively, then a simple proof results from [10]; cf. [12], pp. 681-682. This reduced formulation would suffice in the applications of the Lemma below.

If the C^m-isometry (15) belongs to (1) and to the case

$$E = g_{11}, \qquad F = g_{12}, \qquad G = g_{22} \tag{16}$$

of (14), let T be called a *C^m-autometry* (of (1) or, since T can be replaced by T^{-1}, of (14)). Clearly, the lemma contains the following corollary:

COROLLARY. *Every C^1-autometry T of a regular C^n-metric is a C^{n+1}-autometry* ($n > 0$ is arbitrarily fixed).

Consider two embeddings, say S and S^*, of one and the same metric (1). If one of these two surfaces, say S^*, is of class C^m for a fixed $m > 1$, then its first fundamental form (1) is a regular C^{m-1}-metric. Since S and S^* have the same first fundamental form, it follows from the corollary that the correspondence, induced between S and S^* by their autometry, is represented by a C^m-transformation (15) of non-vanishing Jacobian. Hence (iv) follows from the theorem italicized at the end of p. 612 in [7], if use is made of the Lemma above.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] Bianchi, *Lezioni di geometria differentiale*, 2nd ed. (1902), vol. 1.

[2] S.-S. Chern, P. Hartman and A. Wintner, "On isothermic coordinates," *Commentarii Mathematici Helvetici*, vol. 28 (1954), pp. 301-309.

[3] P. Hartman, "On unsmooth two-dimensional Riemannian metrics," *American Journal of Mathematics*, vol. 74 (1952), pp. 215-226 (Part II, pp. 222-226).

[4] ——— and A. Wintner, "On hyperbolic partial differential equations," *ibid.*, vol. 74 (1952), pp. 834-864.

[5] ——— and A. Wintner, "On the third fundamental form of a surface," *ibid.*, vol. 75 (1953), pp. 298-334.

[6] ——— and A. Wintner, "On pieces of convex surfaces," *ibid.*, vol. 75 (1953), pp. 477-487.

[7] ——— and A. Wintner, "On elliptic Monge-Ampère equations," *ibid.*, vol. 75 (1953), pp. 611-620.

[8] L. Lichtenstein, "Über einige Integrabilitätsbedingungen zweigliedriger Differentialausdrücke mit einer Anwendung auf den Cauchyschen Integralsatz," *Sitzungsberichte der Berliner Mathematischen Gesellschaft*, vol. 9 (1910), pp. 84-100.

[9] H. Weyl, "Über die Bestimmung einer geschlossenen konvexen Fläche durch ihr Linienelement," *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, vol. 61 (1915), pp. 40-72.

[10] A. Wintner, "On Riemann metrics of constant curvature," *American Journal of Mathematics*, vol. 73 (1951), pp. 569-575.

[11] ———, "On isometric surfaces," *ibid.*, vol. 74 (1952), pp. 198-214.

[12] ———, "On the local rôle of the theory of the logarithmic potential in differential geometry," *ibid.*, vol. 75 (1953), pp. 679-690.

ON INDEFINITE BINARY RIEMANNIAN METRICS.*

By AUREL WINTNER.

1. The following considerations were suggested by a theorem which, in the nomenclature used in [14], can be formulated as follows:

(i_n) *If a binary, definite C^n-metric*

$$g_{ik}(u^1, u^2)\,du^i du^k \qquad (g_{12} = g_{21}) \tag{1}$$

is a regular C^n-metric, then it is isometric to a C^n-metric having the conformal normal form

$$\lambda(u, v)\,(du^2 + dv^2) \qquad (\lambda \neq 0) \tag{2}$$

(*and* (2) *is, of course, a regular C^n-metric*).

Here and in the sequel, all assertions are of a local nature, that is, the (u^1, u^2)- or (u, v)-domain is meant to be sufficiently small (in the sense specified at the beginning of [14]).

(i_n) was proved for $n > 1$ in [13], where the limiting case $n = 1$ was left undecided. It was proved in [2] that (i_n) can be extended to this limiting case (which is more intricate than are the cases $n = 2, 3, \cdots$).

In view of the comments made in [13], pp. 686-687, the case $n = 2$ of (i_n) contains a justification of a tacit assumption made by H. Hopf in connection with his elegant formulation of the Codazzi equations (cf. [6], p. 240, formula (4), and p. 234, line 5).

In order to deal with parametrizations of a surface in terms of its lines of curvature and, when the Gaussian curvature is negative, of its asymptotic lines, there is needed an analogue of (i_n) for the case in which the given metric (1) is indefinite; so that what formally corresponds to the normal form (2) is

$$2\lambda(u, v)\,du\,dv \qquad (\lambda \neq 0) \tag{3}$$

(in this connection, cf. [11], pp. 204-205). It will be shown in Sections 7-12 that both parametrization problems, just mentioned, can actually be dealt with on the basis of the following theorem:

(ii_n) *If a binary, indefinite C^n-metric* (1) *is a regular C^n-metric, then*

---

* Received July 27, 1955.

it is isometric to a C^n-metric having the "non-euclidean conformal" normal form (3) (*and* (3) *is, of course, a regular C^n-metric*).

(i_n) deals with a problem in partial differential equations of elliptic type. Correspondingly, the proof of (i_n) depended on potential-theoretical arguments. No such analytic reinforcement can be forthcoming for the problem in (ii_n), since the underlying partial differential equation becomes of hyperbolic type. Thus it is quite unexpected that the counterpart (ii_n) of (i_n) turns out to be true.

What is more, the example given in [4] seemed to indicate that in the applications to theory of surfaces, mentioned above, *two* degrees of differentiability must be sacrificed before the normal form of the parametrization can in general be reached. That actually only *one* degree is lost, will turn out to be a consequence of (ii_n) and of the fact that, in the applications in question, the regularity assumption of (ii_n) happens to be satisfied; cf. the Appendix to [13].

Corresponding to the restriction $n > 1$ on (i_n) in [13], the proof of (ii_n) will be given under the restriction $n > 1$. The limiting index $n = 1$, dealt with in the case of (i_n) in [2], leads in the case of (ii_n) to the same technical complications as in the case of (i_n); complications which undoubtedly can be overcome along the lines of considerations carried out in [2] in the elliptic case and in [5], pp. 357-361, in a similar but hyperbolic (Chebyshev parametrizations).

2. The proof of (ii_n) will be given for $n = 2$. The proof will be such as to make clear its validity for every higher n-value.

Let D be a simply connected, open (u^1, u^2)-domain; it can be assumed to be a sufficiently small neighborhood of the point $(u^1, u^2) = (0, 0)$. The assumptions of (ii_2) are that, on D, the coefficients of (1) are functions of class C^2; that $\det g_{ik} < 0$; finally, that the curvature $K = K(u^1, u^2)$ of (1) is a function of class C^1 (rather than just of class $C^{n-2} = C^{2-2}$, i.e., just continuous; in fact, the regularity of the C^2-metric (1) means, by definition, this increase in the degree of differentiability of K).

For the present, only the C^2-character of (1) will be used (so that $K(u^1, u^2)$ can, for the present, be just continuous). Then, since $\det g_{ik}(u^1, u^2) < 0$ on D, it is clear that the condition

$$g_{ik}(u^1, u^2)\, du^i du^k = 0 \tag{4}$$

can be split into a system of the form

$$du^1/du^2 = f(u^1, u^2), \qquad du^2/du^1 = g(u^1, u^2) \tag{5}$$

in which f and g are real-valued functions of class C^2. Hence (5) has a pair of (local) first integrals, say

$$(6) \qquad u(u^1, u^2) = \text{const.}, \qquad v(u^1, u^2) = \text{Const.},$$

where the two functions u, v of (u^1, u^2) can be chosen to be of class C^2 and of non-vanishing Jacobian. Hence, if T denotes the (local) C^2-transformation $(u^1, u^2) \rightarrow (u, v)$ which results if const. and Const. in (6) are denoted by u and v, respectively, then T^{-1} exists and is of class C^2. Since (4) is equivalent to the system (5), it is clear from the definition of T that T transforms (1) into the normal form (3), and that

$$(7) \qquad \lambda(u, v) \varepsilon C^1 \qquad (\lambda \neq 0)$$

in (3); in fact, the transformation T (along with its inverse T^{-1}) and the mertic (1) are of class C^2. The parenthetical assertion of (7) follows, of course, from $\det g_{ik} \neq 0$.

By using the assumption that (1) is a *regular* C^2-metric, an assumption which was not used thus far, it will be shown in Sections 3-5 that (7) can be improved to

$$(8) \qquad \lambda(u, v) \varepsilon C^2 \qquad (\lambda \neq 0),$$

provided that the C^2-transformation T is suitably chosen. Since (8) means that (3) is a C^2-metric, this will prove (ii$_2$). In fact, the parenthetical assertion made after (3) is clear from the invariance of the scalar $K(u^1, u^2) \varepsilon C^1$ under the mapping T.

3. According to the result (7) of Section 2, the regular C^2-metric (1) is C^2-isometric to a regular C^1-metric (3). If $K(u, v)$ denotes the curvature of (3), then

$$(9) \qquad K(u, v) \varepsilon C^1$$

(by invariance; cf. the end of Section 2). But since (8) is not known, the classical representation of the curvature of (3), the representation supplied by the formula

$$(10) \qquad (\log \lambda)_{uv} = -K\lambda,$$

where $_z = \partial / \partial z$ (cf., e.g., [1], formula (IV) on p. 52), cannot be applied. What can be applied to the C^1-parametrization (3) of the C^2-metric (1) is the integrated form of (10), that is, Weyl's definition ([10], pp. 42-44) of the curvature of a regular C^1-metric; cf., e.g., formula (15) in [13], p. 682 (cf. also the theorem (II) in [5], p. 348, which deals with a more general situation). In the case of the indefinite metric (3), that "integrated form" of the definition of K reduces to

$$(11)\qquad \int_J \{(\log\lambda)_v\,dv - (\log\lambda)_u\,du\} = \int\int_B -2K\lambda\,dudv$$

(cf. (10) or the formula just quoted). Here J denotes any positively oriented, piecewise smooth Jordan curve which, along with its interior domain $B = B(J)$ which occurs on the right of (11), is contained in the (u,v)-domain under consideration.

Choose J to be the boundary of the rectangle B having two points (u,v), $(u+\Delta u, v+\Delta v)$ as opposite vertices. Then the line integral in (11) is readily seen to reduce to $2\Box\log\lambda$, if $\Box\phi$ is defined by

$$(12)\qquad \Box\phi = \phi(u+\Delta u, v+\Delta v) - \phi(u, v+\Delta v) - \phi(u+\Delta u, v) + \phi(u,v).$$

Hence (11) is equivalent to the following identity in (u,v) and $(\Delta u, \Delta v)$:

$$(13)\qquad \Box\log\lambda = \int_u^{u+\Delta u}\int_v^{v+\Delta v} -K\lambda\,dudv.$$

Since K and λ are continuous, (13) implies that, as $(\Delta u, \Delta v) \to (0,0)$, the limit relation

$$(14)\qquad (\Box\log\lambda)/(\Delta u\cdot\Delta v) \to -K(u,v)\lambda(u,v)$$

holds at every point (u,v) of the (u,v)-domain under consideration.

Actually, it is clear from (7), (9), (13) and (12) that $(\log\lambda)_{uv}$ exists and satisfies (10) (so that, in particular, $\lambda_{uv} = \lambda_{vu}$ exists and is continuous). What remains to be shown is that, if the (u,v)-parametrization is suitably chosen, then (10) and (9) improve (7) to (8).

4. To this end, define a function γ of (u,v) by placing

$$(15)\qquad \gamma(u,v) = \int^u\int^v -K\lambda\,dudv,$$

where the choice of the lower limits of integration will be immaterial. It is seen from (7), (9) and (15) that

$$(16)\qquad \gamma(u,v)\ \varepsilon\ C^2.$$

On the other hand, it is clear from (15) that $\gamma_{uv} = -K\lambda$. It follows therefore from (10) that the difference

$$(17)\qquad \phi = \log\lambda - \gamma$$

satisfies the relation

$$\phi_{uv} = 0. \tag{18}$$

Since (18) means that $\phi = \phi(u,v)$ is of the form $\alpha(u) + \beta(v)$, it follows from (17) that

$$\log \lambda(u,v) = \gamma(u,v) + \alpha(u) + \beta(v). \tag{19}$$

Hence, from (7) and (16),

$$\alpha(u) \varepsilon C^1, \qquad \beta(v) \varepsilon C^1. \tag{20}$$

If (19) and (20) are compared with (7) and (16), it follows again (cf. the end of Section 3) that

$$\lambda_{uv}(u,v) = \lambda_{vu}(u,v) \text{ exists and is continuous.} \tag{21}$$

5. Since (10) is a partial differential equation of hyperbolic type, (7) and (21) together contain much less than (8), the assertion to be proved. But it will now be shown that the goal (8) can be reached by an adaptation of a device repeatedly used in [12].

Let the notation be changed in such a way that $\log \gamma$, $\log \alpha$, $\log \beta$ become the functions which in (19) are denoted by γ, α, β, respectively. Then (16) and (20) hold for the new functions also, the new functions are positive, and (3) goes over into

$$2\gamma(u,v)\alpha(u)\beta(v)\,du dv, \tag{22}$$

since (19) becomes $\lambda = \gamma\alpha\beta$.

Next, since $\alpha(u)$, $\beta(v)$ are positive and satisfy (20), it is clear that

$$u' = \int^{u} \alpha(u)\,du, \qquad v' = \int^{v} \beta(v)\,dv \tag{23}$$

defines a C^2-transformation $(u,v) \rightarrow (u',v')$ of non-vanishing Jacobian. Hence, if a function λ' of (u',v') be defined by placing $\lambda'(u',v') = \gamma(u,v)$ (by virtue of (23)), then (16) implies that

$$\lambda'(u',v') \varepsilon C^2, \tag{24}$$

whilst (23) shows that (22) is identical with

$$2\lambda'(u',v')\,du'dv' \qquad (\lambda' > 0). \tag{25}$$

Finally, if the primes are omitted, then (25) and (24) go over into (3) and (8), respectively. The proof is therefore complete.

6. Applications of (ii_n) to the classical parametrization problems of surfaces

(26) $$S\colon X = X(u,v), \text{ where } X = (x,y,z),$$

will now be given. The following terminology will be used:

By a surface S will be meant a sufficiently small piece of a surface, that is, some neighborhood of any given point of the "surface." If S has a parametrization (26) in which the vector $X(u,v)$ is, for some $m \geqq 1$, a function of class C^m on a (sufficiently small) (u,v)-domain D, and if the vector product of the partial derivatives X_u, X_v does not vanish on D, then (26) will be called a C^m-parametrization of S, and S will be referred to as a surface of class C^m (in symbols: $S \varepsilon C^m$).

Note that $k < m$ is quite possible if (26) is a C^k-parametrization of an $S \varepsilon C^m$ (even if the best possible value of k is chosen for the *given* parametrization (21) of $S \varepsilon C^m$). It is precisely this possibility which creates the issues treated below.

7. The following theorem will first be proved:

(iii_m) *If the Gaussian curvature is negative on an $S \varepsilon C^m$, then S possesses C^{m-1}-parametrizations* (26) *in which the parameter lines*

(27) $$u = \text{const.}, \qquad v = \text{Const.}$$

are asymptotic lines of S.

The proof of (iii_m) will depend on the case $n = m - 2$ of (ii_n). Hence (iii_m) will follow for $m > 3$, since (ii_n) was proved for $n > 1$. Correspondingly, (iii_3) is true if (i_1) is. On the other hand, (ii_2) is false (in every sense); cf. [4]. This failure will be seen to be due to the fact that nothing like the result of the Appendix of [13] on an $S \varepsilon C^m$, where $m > 2$, can even be formulated if $m = 2$.

It is an immediate consequence of the case $n = m - 1$ of (i_n) that every $S \varepsilon C^m$ has an isothermic C^m-parametrization; cf. [13], pp. 686-687. In contrast, the "non-euclidean isothermic" parametrization, dealt with by (iii_m), replaces the given C^m-parametrization of S by a C^{m-1}-parametrization, and so one might expect that (iii_m) is not of a final nature. But such is not the case, since the assertion of (iii_m) is as sharp as it can be:

(iii_m bis) *If an S possesses a C^{m-1}-parametrization* (26) *in which the parameter lines* (27) *are asymptotic lines, then $S \varepsilon C^m$.*

In fact, this converse, (iii_m bis), of (iii_m) is known to follow, for every

$m > 1$, from Codazzi's equations (if $m > 2$, and from the "integrated form" of Codazzi's equations if $m = 2$); cf. [3], p. 139.

8. In order to prove (iii$_m$), consider, for a fixed $m > 1$, an $S \varepsilon C^m$ on which the Gaussian curvature does not vanish, and let (1) be the second fundamental form of S with reference to a C^m-parametrization,

$$S\colon\; X = X(u^1, u^2), \tag{28}$$

of S. Then it is clear that (1) is a C^{m-2}-metric of non-vanishing $\det g_{ik}$. Let $m > 2$. Then, according to the Appendix of [13], the C^{m-2}-metric (1) is a regular C^{m-2}-metric. In addition, this metric is indefinite if the Gaussian curvature of S is negative. Hence (ii$_{m-2}$) is applicable.

Accordingly, the parametrization of $S \varepsilon C^m$ can be chosen in such a way that the second fundamental form becomes a regular C^{m-2}-metric having the normal form (3). But the parameter lines (27) are asymptotic lines of S precisely when the second fundamental form is of the form (3). Hence, in order to conclude (iii$_m$), it is sufficient to appeal to a known general theorem, for which various proofs are available (cf. [13], pp. 681-682, where further references are given, and the finer result contained in (II) of [5], p. 348).*

9. In what follows, the problem of those surface parametrizations (26) will be treated in which the parameter lines (27) are required to be lines of curvature. No restriction will now be placed on the sign of the Gaussian curvature (the latter will even be allowed to change sign on S) but umbilical points will, of course, be excluded from S.

Let (26) be a C^m-parametrization of an $S \varepsilon C^m$ and let $m > 1$. Then there exist a first fundamental form

$$E(u,v)\,du^2 + 2F(u,v)\,du\,dv + G(u,v)\,dv^2 \tag{29}$$

and a second fundamental form

$$L(u,v)\,du^2 + 2M(u,v)\,du\,dv + N(u,v)\,dv^2, \tag{30}$$

---

* On p. 360 of the paper [5] of Dr. Hartman and myself, the passage "Hence, . . . above." is erroneous but can easily be rectified by correcting the first four lines as follows:

"Hence $dx + igdy = e^{-i\phi}\omega$ is regular. If (ix) is applied to ω and $R = e^{-i\phi}$ in (u, v)-coordinates, it follows from the choice $g(R) = iR \log R$ that $\phi e^{-i\phi}\omega$ is regular. Since $\phi e^{-i\phi}\omega = \phi(dx + igdy)$, the real forms ϕdx, ϕgdy are regular. Consequently, (vii) implies that ϕ_y and $(\phi g)_x$ exist and are continuous. Thus $g > 0$ and the existence and continuity of g_x imply that ϕ is of class C^1."

In line 9, p. 357, of [5], the relation $\gamma = \rho^{-1}$ should be read as follows: $\gamma = \rho^{-1}\epsilon$.

and the coefficients of (29) and (30) are functions of class C^{m-1} and C^{m-2}, respectively (in terms of the given C^m-parametrization (26) of S).

From the coefficients of (29) and (30), the "Dupin form," say

$$A(u,v)\,du^2+2B(u,v)\,dudv+C(u,v)\,dv^2, \tag{31}$$

can be obtained, by placing (31) equal to

$$\frac{1}{(EG-F^2)^{\frac{1}{2}}}\begin{vmatrix} dv^2 & -dudv & du^2 \\ E & F & G \\ L & M & N \end{vmatrix} \tag{32}$$

(the factor $1/(EG-F^2)^{\frac{1}{2}}$ is attached to the determinant in (32) in order to make the coefficient matrix of (31) a tensor). At the point (u,v) of S, the Dupin form (32) is identical with the geodesic torsion in the direction (du,dv). The coefficients of (31)-(32) are functions of class C^{m-2} in (u,v), since those of (29) and (30) are.

The definition of the lines of curvature of a surface S shows that the parameter lines (27) of a parametrization (26) of S will be lines of curvature if and only if the coefficients of du^2 and dv^2 in (31)-(32) vanish for all (u,v). (Cf., e.g., [1], pp. 98-99). On the other hand, if (26) is *any* parametrization of S (so that $A\equiv 0$ and $B\equiv 0$ are not assumed), then, since the quadratic form (29) is positive definite, the quadratic form (32) is indefinite at all those points (u,v) at which the proportions

$$L:M:N=E:F:G \tag{33}$$

do not hold. But (33) is equivalent to the definition of an umbilical point (u,v) (provided that a "flat" point (u,v), a point at which all three coefficients of (30) vanish, is considered to be an umbilical point). Hence it is clear that if an S is free of umbilical points, then (ii$_n$) supplies conditions which assure the existence of certain parametrizations in which the parameter lines are lines of curvature. What is missing in this plan is the result which, in the application of (ii$_n$) to the problem of asymptotic parametrizations (Section 8), was supplied by the Appendix of [13]. But this will be settled in Sections 11-12.

10. What follows in this manner is the following counterpart (iv$_m$) of (iii$_m$):

(iv$_m$) *If there are no umbilical points on an $S\varepsilon C^m$, then S possesses C^{m-1}-parametrizations* (26) *in which the parameter lines* (27) *are lines of curvature of S.*

Since the proof of (iv_m) depends on the case $n = m - 2$ of (ii_n), the situation with regard to the m-values which are allowed in (iv_m) is exactly the same as specified in the comments which follow the wording of (iii_m) in Section 7 above.

The proof of (iv_m) proceeds as follows: If an $S \varepsilon C^m$, where $m > 1$, is free of umbilical points, then, in terms of a C^m-parametrization of S, the Dupin form of S is an indefinite C^{m-2}-metric (Section 9). It will be shown in Sections 11-12 that, if $m > 2$, then this C^{m-2}-metric is a regular C^{m-2}-metric. Hence it is clear from Section 9 that (iv_m) follows from the case $n = m - 2$ of (ii_n) in the same way as (iii_m) did.

11. Let an $S \varepsilon C^m$, where $m > 2$, be free of umbilical points. Then (32) is a C^{m-2}-metric of non-vanishing determinant (in terms of any C^m-parametrization of S) and has therefore a curvature which, if $m > 3$, is a function of class C^{m-4} (in terms of parameters (u, v) which supply a C^m-parametrization of S). The fact announced in the last paragraph of Section 10 is that this function of class C^{m-4} happens to be of class C^{m-3}, if $m > 2$. This is what will now be verified.

The explicit calculations (which, although they have their geometrical meaning, cannot be avoided) can be simplified by using an isothermic parametrization of S; that is, a parametrization (26) in which (29) reduces to

$$E(u, v)(du^2 + dv^2) \qquad (G = E, F = 0). \tag{34}$$

This assumption involves no loss in the degree of differentiability since, according to [13], p. 687, every $S \varepsilon C^m$ has a C^m-parametrization (26) satisfying $G = E$ and $F = 0$. Then the Dupin form (32) simplifies to

$$M\,du^2 + (N - L)\,du\,dv - M\,dv^2. \tag{35}$$

The standard representations of the mean and Gaussian curvature, $H = H(u, v)$ and $K = K(u, v)$, of S reduce to

$$(36) \quad H = \tfrac{1}{2}(L + N)/E, \qquad (37) \quad K = (LN - M^2)/E^2$$

(cf., e.g., [1], p. 104, where H is the present $-2H$). Finally, as observed by H. Hopf [6], p. 240, the Codazzi equations can be written in the form of an inhomogeneous Cauchy-Riemann system, as follows:

$$M_v - \tfrac{1}{2}(N - L)_u = EH_u, \qquad \tfrac{1}{2}(L - N)_v - M_u = -EH_v. \tag{38}$$

As in the Appendix of [13], suppose first that $m = 4$. Then the assertion is that, if points (u, v) satisfying (33) are excluded, the (indefinite) C^2-metric

(35) has a curvature which is a function of class C^1 in (u, v). This is what will be verified in Section 12. The verification will be such as to make it clear that the proof applies not only for $m = 4$ but also for $m = 5, 6, \cdots$ and, if Weyl's "integrated form" for the curvature of a regular C^1-metric is applied (cf. [10], pp. 42-44), for $m = 3$ also (cf. the corresponding situation in the Appendix of [13]).

12. First, since (26) is a C^4-parametrization,

$$(39) \qquad E(u, v) \varepsilon C^3 \qquad (E > 0)$$

holds for the first fundamental form, whereas the coefficients of (30), and therefore those of (35), are functions of class C^2. According to (36)-(37),

$$(40) \quad H(u, v) \varepsilon C^2, \qquad (41) \quad K(u, v) \varepsilon C^2.$$

Next, if (31) is any C^2-metric, then the explicit form of the prepresentation of its curvature (cf., e. g., [1], p. 51) shows that the principal part of the curvature, that is, the part in which only the contributions of the second derivatives of the coefficients of (31) are retained, is

$$(41) \qquad (A_v - B_u)_v + (C_u - B_v)_u.$$

On the other hand, if (31) is identified with (35) and if only the contribution is retained which, in the preceding sense, corresponds to the "principal part" by virtue of (39), then (41) can be reduced to

$$(42) \qquad (M_v - \tfrac{1}{2}(N - L)_u)_v + (-M_u - \tfrac{1}{2}(N - L)_v)_u.$$

Accordingly, the curvature of (35) will be proved to be of class C^1 if it is ascertained that the function (42) is of class C^1. But (38) shows that (42) differs from

$$(43) \qquad (H_u)_v - (H_v)_u = H_{uv} - H_{vu} = 0$$

only in "non-principal" terms (cf. (39)); so that the proof is complete.

13. If not only the "principal part," but all terms, of the curvature of a C^2-metric (31) are taken into account, then the details of the calculation lead, via (36)-(37) and (38), to the result that the curvature of the case (35) of (31) is

$$(44) \qquad (H^2 - K)^{-2} \partial(H, K)/\partial(u, v)$$

to a trivial factor. The denominator of (44) does not vanish, since (36)-(37) and (34) imply that $H^2 > K$ fails to hold only in the excluded case (33). Clearly, (40)-(41) and (39) imply that (44) is a function of class C^1.

The result (44) can be verified to be equivalent to a classical formula of Weingarten, used by Lie; cf. [1], pp. 234-235. But since Weingarten's proof assumes that the parametrization is in terms of lines of curvature, it leads to a considerable loss in the necessary degree of differentiability. This is the reason why a direct proof had to be given above.

Remark. It is clear that if S has no umbilical points (i.e., if the denominator of (44) does not vanish), then (44) can be a function of class C^{m-3} not only if (26) is a C^m-parametrization (in which case $H \varepsilon C^{m-2}$ and $K \varepsilon C^{m-2}$). In fact, it is sufficient that (26) be a C^{m-1}-parametrization and that the functions H, K, which are then of class C^{m-3}, be such that their Jacobian

$$(45) \qquad \partial(H,K)/\partial(u,v)$$

happens to be a function of class C^{m-2}. Hence, the assumptions of (iv_m) can be generalized so as to involve just the latter conditions.

14. As an application of (iv_m), parametrizations in terms of the particular parameters

$$(46) \qquad u^* = X \cdot N, \qquad v^* = X \cdot X, \quad \text{where } X = (x,y,z),$$

will now be dealt with. Here $N = N(u,v)$ is the unit vector of the oriented normal to S; so that $N(u,v) \varepsilon C^{m-1}$, if (26) is of class C^m. The parameters (u^*, v^*) on S, defined by (46), are intrinsic from the point of view of euclidean geometry, if the indeterminancy of a translation or a rotation of S in the (x,y,z)-space is disregarded.

The replacement of arbitrary parameters (u,v) by the intrinsic parameters (u^*, v^*) first occurs in a note of Weingarten on minimal surfaces [9] (not listed in [1], p. 541). In what follows, Weingarten's somewhat unsymmetric formulae will be replaced by a more straightforward result, and its deduction will be such as to make clear the nature of the conditions to be imposed on S.

It turns out that it is necessary to make the following two assumptions: S is free of umbilical points (i.e.,

$$(47) \qquad H^2 \neq K$$

on S) and S does not have a certain "special" position in the X-space, where $X = (x,y,z)$; cf. (55) below. Note that, whereas (47) is an invariant assumption, the assumption made after (47) refers to the choice of the origin and of the direction of the coordinate axes x, y, z. Such a choice will always be possible, since S is supposed to be a sufficiently small piece of a surface. (It may be mentioned that the assumptions thus imposed on the position of the surface in the X-space, assumptions expressed by (55) below, correspond to what, in connection with the "median surface" of two surfaces, is a "general position" of the latter; cf. [7], p. 757.)

Finally, (iv$_m$) will be needed in order to specify the conditions under which the classical equations defining lines of curvature, that is, the equations

$$N_u = -\alpha X_u, \qquad N_v = -\beta X_v \tag{48}$$

(Rodrigues; cf., e.g., [1], pp. 97-98) are valid with a specified degree of differentiability. In (48), the scalars $\alpha = \alpha(u,v)$, $\beta = \beta(u,v)$ denote the principal curvatures of S; so that

$$H = \tfrac{1}{2}(\alpha + \beta), \qquad K = \alpha\beta. \tag{49}$$

15. It will be concluded from (iv$_m$) that, if an $S \varepsilon C^m$ satisfies the conditions mentioned in Section 14, then S possesses a C^{m-1}-parametrization

$$X = X(u^*, v^*) \tag{50}$$

in terms of the "euclidean" parameters (46) of Weingarten. Incidentally, Weingarten used $v^* = |X|$, instead of the $v^* = |X|^2$ of (46); cf. [9] and the rediscovery of (46) in [8] (and, concerning certain results of [8], the comments in [3], pp. 128-129).

First, since $S \varepsilon C^m$, it follows from (iv$_m$) and (47) that S possesses a C^{m-1}-parametrization (26) in which (48) is valid. Next, since the normal N is attached to $S \varepsilon C^m$ in a parametrization-invariant manner,

$$N(u,v) \varepsilon C^{m-1} \tag{51}$$

(even though only

$$X(u,v) \varepsilon C^{m-1} \tag{52}$$

holds; from (52) alone, only $N(u,v) \varepsilon C^{m-2}$, rather than (51), would follow).

It is clear from (51), (52) and (46) that both functions

$$u^* = u^*(u,v), \qquad v^* = v^*(u,v) \tag{53}$$

are of class C^{m-1}. Hence, if the Jacobian of (53) does not vanish, then the substitution (53) has an inverse of class C^{m-1}. On the other hand, (50) results if the inverse of (53) is substituted into $X(u,v)$. It follows therefore from (52) that the C^{m-1}-character of (50) will be proved if it is ascertained that the Jacobian of (53) does not vanish.

In view of (46) and (48), the Jacobian matrix of (53) is

$$(54)\qquad \begin{pmatrix} u^*_u & u^*_v \\ v^*_u & v^*_v \end{pmatrix} = \begin{pmatrix} -\alpha X\cdot X_u & -\beta X\cdot X_v \\ 2X\cdot X_u & 2X\cdot X_v \end{pmatrix},$$

since the scalar products $N\cdot X_u$, $N\cdot X_v$ vanish identically (N is normal to (26)). But (49) and (47) imply that $\alpha\neq\beta$. Hence the determinant of (54) does not vanish if

$$(55)\qquad X\cdot X_u\neq 0, \qquad X\cdot X_v\neq 0.$$

The pair of conditions (55) is the assumption indicated after (47). Clearly, (55) will be satisfied at every point of any given S if the origin and the direction of the coordinate axes x, y, z in the X-space are chosen in a "non-special" manner (provided that S consists of a sufficiently small neighborhood of any given point of a surface).

16. A parametrization (26) closely related to (50) is

$$(56)\qquad S\colon\ X = X(\alpha, \beta),$$

where α, β are the principal curvatures. (56) is substantially equivalent to the parametrization

$$(57)\qquad S\colon\ X = X(H; K),$$

provided that umbilical points are excluded from S (i.e., that the sign of equality never holds in the inequality

$$(58)\qquad H^2 - K \geqq 0,$$

an inequality which, in view of (49), is always true). In fact, (49) shows that the transition from (56) to (57) is analytic and has therefore an analytic inverse (locally) provided that the Jacobian of (49) does not vanish. But the latter proviso is equivalent to $\alpha\neq\beta$, i.e., to

$$(59)\qquad H^2 \neq K, \qquad\qquad (\text{hence } H^2 > K),$$

since $\alpha = \beta$ is equivalent to

$$(59_0)\qquad H^2 = K.$$

On the other hand, the situation with regard to (57) is as follows:

Suppose that $S \varepsilon C^m$, where $m > 2$, and let (26) be a C^m-parametrization of S. Then

$$(60)\qquad H = H(u, v), \quad K = K(u, v) \text{ are of class } C^{m-2} \qquad (m > 2).$$

Hence, an $S \varepsilon C^m$ possesses a C^{m-2}-parametrization (57) whenever the Jacobian (45) does not vanish. In fact, the substitution (60) then possesses an inverse, $(H, K) \rightarrow (u, v)$, which is of class C^{m-2}, and (26) is identical with (57) by virtue of this substitution.

It seems to be peculiar that only

$$(61) \qquad \partial(H,K)/\partial(u,v) \neq 0,$$

rather than (61) *and* (59), was used in this deduction. The explanation is that (59_0) cannot hold unless

$$(61_0) \qquad \partial(H,K)/\partial(u,v) = 0.$$

17. Let a point (u, v) of an $S \varepsilon C^3$ be called a *W-point* if, in some (but then, of course, in every) C^3-parametrization (26) of S, the Jacobian (45) vanishes at that point. By covariance, the same must then hold if (26) is any C^1-parametrization (for instance, any C^2-parametrization) of $S \varepsilon C^3$. It is also clear that an $S \varepsilon C^3$ is what in a customary sense of the word is a *W-surface* (Weingarten) if and only if every point of S is a W-point.

Every umbilical point of an $S \varepsilon C^3$ is a W-point. In fact, a comparison of (58) with (59_0) shows that, if (u, v) is an umbilical point, then the function $H^2 - K$ acquires a (local, not necessarily isolated) minimum and has, therefore, a stationary value. But then the gradient of $H^2 - K$ vanishes. This means that the vector $(2H, -1)$ is a solution of two homogeneous, linear equations having the determinant (45), which proves that (61_0) is satisfied.

The converse of this deduction is not possible, i. e., a W-point need not be a stationary point of the function $H^2 - K$. In fact, the latter is known to be stationary at every point of (that is, constant on) an $S \varepsilon C^3$ only if S is a plane, a sphere or a circular cylinder, whereas (45) will vanish identically on every surface of revolution.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] L. Bianchi, *Lezioni di geometria differenziale*, 1st ed. (1894).

[2] S.-S. Chern, P. Hartman and A. Wintner, "On isothermic coordinates," *Commentarii Mathematici Helvetici*, vol. 28 (1954), pp. 301-309.

[3] P. Hartman and A. Wintner, "On the curvatures of a surface," *American Journal of Mathematics*, vol. 75 (1953), pp. 127-141.

[4] ——— and A. Wintner, "On asymptotic parametrizations," *ibid.*, vol. 75 (1953), pp. 486-496.

[5] ——— and A. Wintner, "Regular binary Pfaffians and non-parabolic partial differential equations," *Rendiconti del Circolo Matematico di Palermo*, ser. 2, vol. 3 (1955), pp. 347-362.

[6] H. Hopf, "Über Flächen mit einer Relation zwischen den Hauptkrümmungen," *Mathematische Nachrichten*, vol. 4 (1951), pp. 232-249.

[7] ——— and H. Samelson, "Zum Beweis des Kongruensatzes für Eiflächen," *Mathematische Zeitschrift*, vol. 43 (1938), pp. 749-766.

[8] W. Scherrer, "Stützfunktion und Radius. I.," *Commentarii Matematici Helvetici*, vol. 20 (1947), pp. 366-381.

[9] J. Weingarten, "Eine neue Klasse auf einander abwickelbarer Flächen," *Göttinger Nachrichten* (1887), pp. 28-31.

[10] H. Weyl, "Ueber die Bestimmung einer geschlossenen konvexen Fläche durch ihr Linienelement," Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich, vol. 61 (1915), pp. 40-72.

[11] A. Wintner, "On isometric surfaces," *American Journal of Mathematics*, vol. 74 (1952), pp. 198-214.

[12] ———, "On the infinitesimal geometry of curves," *ibid.*, vol. 75 (1953), pp. 241-259.

[13] ———, "On the local rôle of the theory of the logarithmic potential in differential geometry," *ibid.*, vol. 75 (1953), pp. 679-690.

[14] ———, "On the local embedding problems in the differential geometry of surfaces," vol. 77 (1955), pp. 845-852.

ON THE GALOIS THEORY OF DIFFERENTIAL FIELDS.*

By E. R. Kolchin.[1]

Introduction.

1. Summary. In a preceding paper [7] there was presented a Galois theory, for a certain kind of differential field extension called strongly normal. The Galois group of a strongly normal extension is endowed with a structure very much like that of a group variety, as studied by Weil [14]. In the present paper the study of such Galois groups is renewed for the purpose of clarifying their connection with group varieties on the one hand and with strongly normal extensions on the other.

Consider a differential field $\mathfrak{F}$ of characteristic 0 with algebraically closed field of constants $\mathcal{C}$, and a strongly normal extension $\mathfrak{G}$ of $\mathfrak{F}$; suppose given a universal extension $\mathfrak{F}^*$ of $\mathfrak{G}$, and denote the field of constants of $\mathfrak{F}^*$ by $\mathcal{C}^*$. For reasons of convenience we use the term "Galois group of $\mathfrak{G}$ over $\mathfrak{F}$" for the group of all (ipso facto strong) isomorphisms of $\mathfrak{G}$ over $\mathfrak{F}$, and not for its subgroup consisting of all automorphisms of $\mathfrak{G}$ over $\mathfrak{F}$. The field $\mathcal{C}^*$, being algebraically closed and of finite transcendence degree over $\mathcal{C}$, may be used as a universal domain for algebraic geometry (Weil [13], ch. I, § 1); it is the group varieties of this algebraic geometry, defined over $\mathcal{C}$ and slightly generalized to permit group varieties which are reducible (i. e. which have more than one component), which we consider. By an "algebraic group" we mean either a Galois group or a group variety as above.

Chapter I is primarily a study of "rational" homomorphisms of algebraic groups into algebraic groups; these seem to be the homomorphisms appropriate to the consideration of algebraic groups. Rational homomorphisms of group varities into group varieties, without restriction to fields of characteristic 0, were considered by Weil [14] (who omitted the adjective "rational").

Specializing the concept of rational homomorphism we obtain that of birational isomorphism. In Chapter II it is shown that every Galois group is birationally isomorphic to a group variety, and conversely that every *irreducible* group variety is birationally isomorphic to a Galois group; this

---

* Received June 29, 1955.

[1] Some of the results in this paper were obtained in connection with a contract with the National Science Foundation, and some while the author was a Guggenheim fellow.

answers a question raised in [7], p. 759. The first of these results has already been proved by Matsumura [8] for irreducible Galois groups (the extension to reducible Galois groups is fairly trivial); the present proof, although shorter, makes use of a special theorem due to Weil and refined by Rosenlicht (see Chapter II, § 1 below), whereas Matsumura's proof has the merit of being more self-contained. The proof of the second of these results makes use of the fact that the vector space of derivations of the field of rational functions on an irreducible group variety has a base consisting of invariant derivations (see Chapter II, § 2 below); this fact, which is central in the proof, was pointed out to me in this connection by Chevalley. Whether or not the result can be extended to reducible group varieties remains open. Chapter II also contains certain consequences of these two results.

In Chapter III the concept of a rational homomorphism of the Galois group $\mathfrak{G}$ of $\mathcal{G}$ over $\mathcal{F}$ into a group variety G is generalized to obtain that of a rational crossed homomorphism, or a cocycle of dimension 1, or a 1-cocycle, of $\mathfrak{G}$ into G. This leads, in much the usual way, to the set $H^1(\mathfrak{G}, G)$ of cohomology classes of dimension 1 of $\mathfrak{G}$ into G. Since G is not assumed to be commutative, $H^1(\mathfrak{G}, G)$ is not in general a group. It is proved that $H^1(\mathfrak{G}, G)$ is trivial provided *either* G is a full matric group *or* $\mathcal{F}$ is algebraically closed; also, if G is commutative (so that $H^1(\mathfrak{G}, G)$ is a group) then every element of $H^1(\mathfrak{G}, G)$ has finite order. Using the first of these results, it is shown that the strongly normal extension $\mathcal{G}$ of $\mathcal{F}$ is a Picard-Vessiot extension if and only if its Galois group $\mathfrak{G}$ is birationally isomorphic to an algebraic group of matrices; this answers another question raised in [7], p. 759. Similarly, the last of these results is used to give a short proof of the characterization given in [7] of all strongly normal extensions of transcendence degree 1.

2. Assumptions and notation. Throughout this paper $\mathcal{F}$ denotes a differential field of characteristic 0 with derivation operators $\delta_1, \cdots, \delta_m$ and algebraically closed field of constants $\mathcal{C}$, and $\mathcal{G}$ denotes a strongly normal extension of $\mathcal{F}$. We fix a universal extension $\mathcal{F}^*$ of $\mathcal{G}$ and denote the field of constants of $\mathcal{F}^*$ by $\mathcal{C}^*$; except when the contrary is explicitly stated, all differential fields considered are differential subfields of $\mathcal{F}^*$, so that every isomorphism of a differential field has image in $\mathcal{F}^*$.

Because $\mathcal{G}$ is strongly normal over $\mathcal{F}$, every isomorphism σ of $\mathcal{G}$ over $\mathcal{F}$ is strong; denoting the field of constants of $\mathcal{G}\langle\sigma\mathcal{G}\rangle$ by $\mathcal{C}(\sigma)$ (instead of by $\mathcal{C}_\sigma$ as in [7]), we can express this fact by writing

$$\mathcal{G}\langle\sigma\mathcal{G}\rangle = \mathcal{G}\langle\mathcal{C}(\sigma)\rangle = (\sigma\mathcal{G})\langle\mathcal{C}(\sigma)\rangle.$$

As in [7], we identify each such σ with the unique automorphism of $\mathcal{G}\langle\mathcal{C}^*\rangle$ over $\mathcal{F}\langle\mathcal{C}^*\rangle$ which extends σ; the set of all isomorphisms of $\mathcal{G}$ over $\mathcal{F}$ is thereby identified with the group of all automorphisms of $\mathcal{G}\langle\mathcal{C}^*\rangle$ over $\mathcal{F}\langle\mathcal{C}^*\rangle$. We call this group the *Galois group* of $\mathcal{G}$ over $\mathcal{F}$ and denote it by $\mathfrak{G}$ (instead of by $\mathfrak{G}^*$ as in [7]); we recall that $\mathfrak{G}$ is algebraic in the sense defined in [7]. The group of all automorphisms of $\mathcal{G}$ over $\mathcal{F}$ is a subgroup of $\mathfrak{G}$. If $\sigma_1, \cdots, \sigma_p \in \mathfrak{G}$ we denote by $\mathcal{C}(\sigma_1, \cdots, \sigma_p)$ the field of constants of $\mathcal{G}\langle\sigma_1\mathcal{G}, \cdots, \sigma_p\mathcal{G}\rangle$, that is, the compositum of the p fields $\mathcal{C}(\sigma_1), \cdots, \mathcal{C}(\sigma_p)$. It is easy to see that $\mathcal{C}(\sigma^{-1}) = \mathcal{C}(\sigma)$ and $\mathcal{C}(\sigma\tau, \sigma) = \mathcal{C}(\sigma, \tau) = \mathcal{C}(\sigma\tau, \tau)$ for all $\sigma, \tau \in \mathfrak{G}$.

Chapter I. Rational homomorphisms.

1. Varieties. The field $\mathcal{C}^*$ is an algebraically closed extension of $\mathcal{C}$ of infinite transcendence degree, and therefore may be used as a universal domain for algebraic geometry. We find it convenient to deviate from Weil's terminology, in order to permit varieties which are not irreducible. Specifically, we adopt the following terminology, which is sufficient for our purpose:[2] we call *irreducible variety* what Weil calls abstract variety; we call *variety* any finite union of irreducible subvarieties of some irreducible variety (in Weil's terminology, a bunch of subvarieties of an abstract variety), or the union of a disjoint finite set of such. Every irreducible variety of dimension > 0 which we consider, except those for which the contrary is explicitly stated, is assumed to have $\mathcal{C}$ as a field of definition. Similarly, when we speak of a specialization of a point it is always to be understood, except when the contrary is stated, that the specialization is over $\mathcal{C}$. We find it convenient sometimes to call a subset of a variety *algebraic* to indicate that the subset is a subvariety. What we mean by the *components* of a variety is clear.

Let V and W be varieties, and denote the components of V by $V_1, \cdots, V_h$.

---

[2] An intrinsic definition of abstract algebraic variety, encompassing the case of reducible varieties, has been given by Serre [12]. The present definition, which involves an imbedding in some irreducible variety, is adopted here in the interest of brevity. This is, of course, merely a difference of procedure and not of substance. However, in the discussion below of rational mappings, the definition adopted for what it means for a rational mapping to be defined at a particular point is weaker than the definition that naturally arises in developing farther the ideas of Serre. Specifically, the present definition leads to a notion of isomorphism of varieties such that two components of a variety can have different higher intersection properties from the corresponding components of an isomorphic variety; on the other hand, because tthe definition here is weaker, Proposition 1 below is stronger. Of course, when dealing with varieties for which the components are disjoint (which by Proposition 1 is the case for group varieties) this difference in the two definitions disappears.

By a *rational mapping* of V into W we mean any sequence $f = (f_1, \cdots, f_h)$ where, for each i, f_i is a rational mapping of V_i into some component of W (that is, f_i is what is called by Weil [14] a function defined on V_i with values in a component of W); by a *field of definition* of f we mean a common field of definition of $f_1, \cdots, f_h$. All rational mappings except those for which the contrary is stated, are assumed to have $\mathcal{E}$ as a field of definition. We say the rational mapping f is *defined* at a point $s \varepsilon V$ if, for every point $t \varepsilon V$ of which s is a specialization, there exists an index i such that $t \varepsilon V_i$ and f_i is defined at t, and for all such indices i the value $f_i(t)$ is the same. If f is defined at s then there exists an i such that f_i is defined at s, and the point $f_i(s)$ is independent of the choice of i; we call this point the *value* of f at s and denote it by $f(s)$. This definition insures that if s' is a specialization of $s \varepsilon V$ and f is defined at s', then f is defined at s and $(s', f(s'))$ is a specialization of $(s, f(s))$. If ϕ is a mapping (in the set-theoretic sense) of a subset of V into W, then we permit ourselves to say that ϕ is *rational* if there exists a rational mapping f of V into W such that f is defined at precisely those points $s \varepsilon V$ at which ϕ is defined and $f(s) = \phi(s)$ at all such points s; in other words, we identify each rational mapping of V into W with a mapping, which we call rational.

By a *biregular birational mapping of V onto W* we mean a one-to-one mapping f of V onto W such that f and f^{-1} are both rational and everywhere defined.

By a *group variety* we mean any variety G with a group structure such that (when the group composition is written multiplicatively) the mapping $(s, s') \rightarrow ss'$ of $G \times G$ into G and the mapping $s \rightarrow s^{-1}$ of G into G are rational. Using an argument which is familiar in several special cases and analogous situations ([5], § 2, Th., p. 10; Weil [14], § 4, Prop. 9, pp. 39-40; Chevalley [3], Ch. I, § 3, Th. 2, p. 86; [7], Ch. II, Th. 1, p. 788), it is easy to prove the following result.

PROPOSITION 1. *Precisely one component of any group variety G contains the unity element; this component G^0 is a normal subgroup of finite index, and the cosets of G relative to G^0 are the components of G.*

Thus, the components of G are disjoint and all have the same dimension, which we call the dimension of G and denote by $\dim G$.

PROPOSITION 2. *If $\mathfrak{h}$ is a subgroup of the group variety G consisting solely of points rational over $\mathcal{E}$, and H is the smallest subvariety of G containing $\mathfrak{h}$, then H is a subgroup of G.*

Proof. For any $s \varepsilon \mathfrak{h}$, sH is the smallest subvariety of G containing $s\mathfrak{h} = \mathfrak{h}$, so that $sH = H$; therefore $\mathfrak{h}H = H$. For every $t \varepsilon H$ which is rational over $\mathcal{K}$, Ht is the smallest subvariety of G containing $\mathfrak{h}t$, and $\mathfrak{h}t \subset \mathfrak{h}H = H$, so that $Ht \subset H$; this inclusion obviously continues to hold without the restriction that t be rational over $\mathcal{K}$, so that $HH \subset H$. Finally, H^{-1} is the smallest subvariety of G containing $\mathfrak{h}^{-1} = \mathfrak{h}$, so that $H^{-1} = H$. Therefore H is a group.

PROPOSITION 3. *If $\mathfrak{g}$ is a subgroup of the group variety G such that $G - \mathfrak{g}$ is contained in a subvariety of G all the components of which are of lower dimension than G, then $\mathfrak{g} = G$.*

Proof. If $\mathfrak{g}$ were different from G, a coset $s\mathfrak{g}$ of G relative to $\mathfrak{g}$ would be contained in $G - \mathfrak{g}$ and therefore in a subvariety of G all the components of which have dimension $< \dim G$, and we obviously may suppose s rational over $\mathcal{K}$; therefore $\mathfrak{g}$ would be contained in such a variety, so that $G = \mathfrak{g} \cup (G - \mathfrak{g})$ would, too, contradiction.

The proofs of Propositions 2 and 3 are the same as those of the analogous propositions on algebraic groups of strong isomorphisms of a differential field (see [7], Ch. II, Props. 14, 15, p. 789, for the case of algebraic groups of automorphisms).

2. Rational homomorphisms of group varieties. Consider two group varieties G and H, and a rational mapping f of G into H such that $f(st) = f(s)f(t)$ whenever s and t are points of G such that f is defined at s, t, and st (or what amounts to the same thing, whenever s and t are independent generic points of components of G). We call such a rational mapping a *rational homomorphism* of G into H.

PROPOSITION 4. *A necessary and sufficient condition that f be a rational homomorphism of the group variety G into the group variety H is that f be a homomorphism of the group G into the group H having the two properties: 1) for all $s \varepsilon G$, $\mathcal{K}(f(s)) \subset \mathcal{K}(s)$; 2) for all $s, s' \varepsilon G$ with s' a specialization of s, $f(s')$ is a specialization of $f(s)$.*

Proof. Let f be a rational homomorphism of G into H, and let s' be any point of G. Choose independent generic points s, t of components of G such that s' is a specialization of s; there exists a specialization t' of t such that t' is rational over $\mathcal{K}$ and f is defined at t' and $s't'$. Then f is defined at st', and $f(s) = f(st')f(t')^{-1}$. This makes it clear that f is defined at s', that is, f is defined at every point of G. Thus, f is a homomorphism (in the group theoretic sense) of the group G into the group H, and f obviously has the two required properties.

Conversely, let f be any homomorphism of the group G into the group H having these two properties. Letting $s_1, \cdots, s_n$ be independent generic points of the n components of G, we see from 1) that there exists a rational mapping g of G into H such that $g(s_i) = f(s_i)$ $(1 \leqq i \leqq h)$; we then see from 2) that $g(s) = f(s)$ for every generic point of every component of G, so that $g(st) = g(s)g(t)$ whenever s and t are independent generic points of components of G; in other words, g is a rational homomorphism of G into H. Therefore g is everywhere defined and, by 2), $g(s) = f(s)$ for every $s \varepsilon G$.

PROPOSITION 5. *Let f be a rational homomorphism of a group variety G into a group variety H. If N is any algebraic set in H then $f^{-1}(N)$ is an algebraic set in G. If M is an algebraic subgroup of G then $f(M)$ is an algebraic subgroup of H.*

Proof. Let A be the smallest algebraic set in G containing $f^{-1}(N)$, let A_0 be any component of A, and let s be a generic point of A_0. If $f(s)$ were not a point of N there would exist a proper algebraic subset A_1 of A_0 such that every specialization s' of s with $s' \not\varepsilon A_1$ has the property that $f(s') \not\varepsilon N$, that is $f^{-1}(N)$ would be contained in the algebraic set obtained from A on replacing the component A_0 by A_1, so that A would not be the smallest alegbraic set in G containing $f^{-1}(N)$. Therefore $f(s) \varepsilon N$, so that $f(A_0) \subset N$, whence $f(A) \subset N$. Thus $f^{-1}(N) = A$, which is algebraic.

Let B be the smallest algebraic set in H containing $f(M)$; obviously B is also the smallest algebraic set in H containing $f(M')$, where M' is the group of all points of M which are rational over $\mathcal{G}$, so that by Proposition 2 B is an algebraic group. Let $s_1, \cdots, s_h$ be generic points of all the components of M, and let B_i be the irreducible algebraic set in H with generic point $f(s_i)$. Clearly $B = B_1 \cup \cdots \cup B_h$. For each i there exists a proper algebraic subset C_i of B_i such that every specialization t_i' of $f(s_i)$ with $t_i' \not\varepsilon C_i$ has the property that $t_i' = f(s_i')$ for some specialization s_i' of s_i. Thus, $B - f(M)$ is contained in an algebraic set every component of which is a proper subset of a component of B. It follows from Proposition 3 that $f(M) = B$.

COROLLARY. *The kernel and the image of a rational homomorphism are algebraic.*

We call a rational homomorphism of a group variety G into a group variety H a *birational isomorphism* of G onto H provided the kernel is trivial, the image is H, and the inverse is a rational homomorphism of H into G. It follows from Proposition 4 that f is a birational isomorphism of G onto H if and only if f is an isomorphism of the group G onto the group H such that: 1) $\mathcal{C}(s) = \mathcal{C}(f(s))$ for every $s \varepsilon G$; 2) for all s, s' in G, s' is a specialization of s if and only if $f(s')$ is a specialization of $f(s)$.

3. Rational homomorphisms in general. Henceforth we use the term "algebraic group" to mean *either* a Galois group (of a strongly normal extension of a differential field with field of constants $\mathcal{C}$), *or else* a group variety. We generalize to arbitrary algebraic groups the notion of rational homomorphism previously defined for group varieties.

By a *rational homomorphism* of an algebraic group G into an algebraic group H we mean any homomorphism f of the group G into the group H having the two properties that whenever $s \varepsilon G$ then $\mathcal{C}(f(s)) \subset \mathcal{C}(s)$, and whenever s, $s' \varepsilon G$ and s' is a specialization of s then $f(s')$ is a specialization of $f(s)$. If in addition the kernel of f is trivial and the image of f is H, so that f is an isomorphism of the group G onto the group H, and if f^{-1} is a rational homomorphism of H into G, then we say that f is a *birational isomorphism* of G onto H. In other words, a birational isomorphism of G onto H is an isomorphism f of the group G onto the group H such that $\mathcal{C}(f(s)) = \mathcal{C}(s)$ $(s \varepsilon G)$ and s' is a specialization of s if and only if $f(s')$ is a specialization $f(s)$ $(s, s' \varepsilon G)$.

Example 1. Let $\mathcal{M}$ be a differential field between $\mathcal{F}$ and $\mathcal{G}$ which is strongly normal over $\mathcal{F}$, and denote the Galois group of $\mathcal{M}$ over $\mathcal{F}$ by $\mathfrak{H}$. We know ([7], Ch. III, Th. 4, p. 797) that the algebraic subgroup $\mathfrak{G}(\mathcal{M})$ of $\mathfrak{G}$ corresponding to the intermediate differential field $\mathcal{M}$ is normal, and that the mapping r which to each $\sigma \varepsilon \mathfrak{G}$ corresponds the restriction of σ to $\mathcal{M}$ is a homomorphism, with kernel $\mathfrak{G}(\mathcal{M})$ and image $\mathfrak{H}$, of the group $\mathfrak{G}$ into the group $\mathfrak{H}$. For every $\sigma \varepsilon \mathfrak{G}$, the field of constants of $\mathcal{M}\langle r(\sigma)\mathcal{M}\rangle = \mathcal{M}\langle\sigma\mathcal{M}\rangle$ is obviously contained in the field of constants of $\mathcal{G}\langle\sigma\mathcal{G}\rangle$, that is, $\mathcal{C}(r(\sigma)) \subset \mathcal{C}(\sigma)$. Also, if $(\sigma'\alpha)_{\alpha \varepsilon \mathcal{G}}$ is a specialization of $(\sigma\alpha)_{\alpha \varepsilon \mathcal{G}}$ over $\mathcal{G}$ then $(r(\sigma')\alpha)_{\alpha \varepsilon \mathcal{H}} = (\sigma'\alpha)_{\alpha \cdot \mathcal{H}}$ is a specialization of $(r(\sigma)\alpha)_{\alpha \varepsilon \mathcal{H}} = (\sigma\alpha)_{\alpha \varepsilon \mathcal{H}}$ over $\mathcal{M}$, that is, if σ' is a specialization of σ then $r(\sigma')$ is a specialization of $r(\sigma)$. Therefore r is a rational homomorphism of $\mathfrak{G}$ into $\mathfrak{H}$.

Example 2. If $\mathcal{G}$ is a Picard-Vessiot extension of $\mathcal{F}$ then ([5] and [6]) $\mathcal{G}$ has a fundamental system of generators, that is, a finite number of elements

$\eta_1, \cdots, \eta_n$ linearly independent over constants such that $\mathcal{G} = \mathcal{F}\langle \eta_1, \cdots, \eta_n \rangle$ and such that for every $\sigma \varepsilon \mathfrak{G}$ there exists an invertible matrix of constants

$$c(\sigma) = (c_{ij}(\sigma))_{1 \leqq i \leqq n,\, 1 \leqq j \leqq n}$$

for which

$$\sigma \eta_j = \sum_{i=1}^{n} c_{ij}(\sigma) \eta_i \qquad (1 \leqq j \leqq n).$$

The mapping $c: \sigma \to c(\sigma)$ $(\sigma \varepsilon \mathfrak{G})$ is an isomorphism of $\mathfrak{G}$ onto an algebraic subgroup of the group variety $GL(n)$ of all invertible square matrices of degree n with coordinates in $\mathcal{K}^*$. It is not difficult to see that c is a birational isomorphism.

It is obvious that if f is a rational homomorphism of an algebraic group G into an algebraic group H and g is a rational homomorphism of H into an algebraic group I, then $g \circ f$ is a rational homomorphism of G into I.

PROPOSITION 6. *Let f and f' be rational homomorphisms of an algebraic group G into group varieties H and H' respectively. If f and f' have the same kernel and have respective images H and H', then the isomorphism ϕ of H onto H' such that $\phi \circ f = f'$ is birational.*

Proof. Letting s be any element of G we see that for a generic specialization s' of s, $f(s) = f(s')$ if and only if $f'(s) = f'(s')$; this means that an isomorphism of $\mathcal{K}(s)$ leaves invariant each element of $\mathcal{K}(f(s))$ if and only if it leaves invariant each element of $\mathcal{K}(f'(s))$, so that $\mathcal{K}(f(s)) = \mathcal{K}(f'(s))$; thus $\mathcal{K}(t) = \mathcal{K}(\phi(t))$ for every $t \varepsilon H$. Now, if s is a generic point of a component of G then $f(s)$ is a generic point of a component of H, and every generic point of a component of H is of this form; similarly for H'. If t' is a generic specialization of $t = f(s)$, then there exists a generic specialization s' of s such that $t' = f(s')$, and obviously $f'(s')$ is a generic specialization of $f'(s)$. Therefore there exists a rational mapping g of H into H' such that $g(t) = \phi(t)$ for every generic point t of a component of H. If t, t' are independent generic points of components of H then we may write $t = f(s)$, $t' = f(s')$, where s, s' are independent generic elements of components of G, so that $g(tt') = g(f(s)f(s')) = g(f(ss'))$ $= \phi(f(ss')) = f'(ss') = f'(s)f'(s') = \phi(t)\phi(t') = g(t)g(t')$. Therefore g is a rational homomorphism of H into H'. It follows that $g \circ f$ is a rational homomorphism of G into H' coinciding with f' at every generic point of every component of G and therefore at every point whatever, that is $g \circ f = f'$. Therefore $\phi = g$, so that ϕ is a rational homomorphism of H into H'.

Reversing the roles of f and f' we see that ϕ^{-1} is rational too, so that ϕ is birational.

COROLLARY. *If f is a rational homomorphism of a group variety G into a group variety H such that the kernel of f is trivial and the image of f is H, then f is a birational isomorphism of G onto H.*

Proof. In the proposition set $H' = G$, $f' =$ the identity mapping of G.

PROPOSITION 7. *Let G be an algebraic group, G_1 a normal algebraic subgroup of G of finite index, and K any normal algebraic subgroup of G. Then every rational homomorphism f_1 of G_1 with kernel $G_1 \cap K$ and image a group variety can be extended to a rational homomorphism f of G with kernel K and image a group variety.*

Proof. We extend f_1 in two stages; first to a rational homomorphism of G_1K with kernel K and the same image as f_1, and then to a rational homomorphism of G. G_1K is the union of its cosets relative to G_1; these are finite in number and each is an algebraic set, so that G_1K is a normal algebraic subgroup of G. Every element s of G_1K is of the form $s = s_0a$, where $s_0 \varepsilon G_1$ and $a \varepsilon K$, and $f_1(s_0)$ clearly depends only on s and not on the particular choice of s_0 and a. Therefore we may define a mapping g of G_1K into the image H_1 of f_1 by the formula $g(s_0a) = f_1(s_0)$. The computation $g(s_0'a's_0a) = g(s_0's_0 \cdot s_0^{-1}a's_0a) = f_1(s_0's_0) = f_1(s_0')f_1(s_0) = g(s_0'a')g(s_0a)$ shows that g is a homomorphism of G_1K with image H_1, and clearly g has kernel K. Now every coset of G_1K relative to G_1 contains an element of K, and therefore an element $a \varepsilon K$ such that $\mathscr{k}(a) = \mathscr{k}$. Therefore every element $s \varepsilon G_1K$ can be written in the form $s = s_0a$ with $s_0 \varepsilon G_1$, $a \varepsilon K$, and $\mathscr{k}(a) = \mathscr{k}$, so that $\mathscr{k}(g(s)) = \mathscr{k}(f(s_0)) \subset \mathscr{k}(s_0) = \mathscr{k}(s_0a) = \mathscr{k}(s)$; furthermore, if s' is a specialization of s then s' belongs to the same coset as s, so that $s' = s_0'a$, where $s_0' \varepsilon G_1$; obviously s_0' is a specialization of s_0, so that $f_1(s_0')$ is a specialization of $f_1(s_0)$, whence $g(s')$ is a specialization of $g(s)$. Thus, the homomorphism g is rational.

Let $s_1, \cdots, s_r$ be a complete set of representatives of the cosets of G relative to G_1K; we assume without loss of generality that $\mathscr{k}(s_i) = \mathscr{k}$ $(1 \leqq i \leqq r)$ and s_1 is the unity element of G. For each index i the mapping ϕ_i of G_1K defined by $\phi_i(s) = s_iss_i^{-1}$ is clearly a birational isomorphism of G_1K onto G_1K. For each pair of indices i, j there is an index $k(i,j)$ and an element $s(i,j) \varepsilon G_1K$ such that $s_is_j = s(i,j)s_{k(i,j)}$. Let $H_2, \cdots, H_r$ be varieties biregularly birationally equivalent to H_1 such that $H_1, H_2, \cdots, H_r$ are disjoint, and let h_i be a biregular birational mapping of H_1 into H_i $(1 \leqq i \leqq r)$, where h_1 is the identity mapping of H_1.

Since g and $g \circ \phi_i$ are rational homomorphisms of G_1K with the same kernel K and image H_1, there exists by Proposition 6 a birational isomorphism ψ_i of H_1 onto H_1 such that $\psi_i \circ g = g \circ \phi_i$. Furthermore, the mapping ρ_{ij} of H_1 defined by $\rho_{ij}(t) = t \cdot g(s(i,j))$ is a biregular birational mapping of H_1 onto H_1. Also, the law of composition, which we denote by μ_1, of the group variety H_1 is an everywhere defined rational mapping of $H_1 \times H_1$ into H_1. Consequently the mapping

$$\mu_{ij} = h_{k(i,j)} \circ \rho_{ij} \circ \mu_1 \circ (h_i^{-1} \times (\psi_i \circ h_j^{-1}))$$

is an everywhere defined rational mapping of $H_i \times H_j$ into the variety $H = H_1 \cup \cdots \cup H_r$, and it is easy to see that $\mu_{11} = \mu_1$. There is a unique everywhere defined rational law of composition μ on H which extends every μ_{ij}; since μ extends μ_1 we may write $\mu(t, t') = tt'$ for $t, t' \varepsilon H$.

Let f be the mapping of G into H defined by $f(ss_i) = h_i(g(s))$ $(s \varepsilon G_1K, 1 \leqq i \leqq r)$. Clearly f has image H, extends f_1, and because

$$\begin{aligned} f(ss_i \cdot s's_j) &= f(s \cdot s_is's_i^{-1} \cdot s_is_j) = f(s\phi_i(s')s(i,j)s_{k(i,j)}) \\ &= h_{k(i,j)}(g(s)g(\phi_i(s'))g(s(i,j))) \\ &= (h_{k(i,j)} \circ \rho_{ij} \circ \mu_1 \circ (h_i^{-1} \times (\psi_i \circ h_j)))(h_i(g(s)), h_j(g(s'))) \\ &= h_i(g(s))h_j(g(s')) = f(ss_i)f(s's_j), \end{aligned}$$

f is a homomorphism of the group G. Therefore H, with the law of composition μ, is a group. To prove that H is a group variety it remains to show that $t \to t^{-1}$ $(t \varepsilon H)$ is an everywhere defined rational mapping of H into H. Now, the mappings λ_i and ρ_i of H defined by $\lambda_i(t) = f(s_i)^{-1}t$ and $\rho_i(t) = tf(s_i)^{-1}$ are everywhere defined rational mappings of H into H mapping H_i into H_1, and the mapping ν of H_1 defined by $\nu(t) = t^{-1}$ is an everywhere defined rational mapping of H_1 into H; but the mapping $t \to t$ $(t \varepsilon H)$ obviously coincides with $\lambda_i \circ \nu \circ \rho_i$ on H_i and therefore is rational and everywhere defined.

Finally, if $s \varepsilon G_1K$ then

$$\mathscr{K}(f(ss_i)) = \mathscr{K}(h_i(g(s))) = \mathscr{K}(g(s)) \subset \mathscr{K}(s) = \mathscr{K}(ss_i),$$

and if $s's_i$ is a specialization of ss_i then s' is a specialization of s, $g(s')$ is a specialization of $g(s)$, and $f(s's_i) = h_i(g(s'))$ is a specialization of $f(ss_i) = h_i(g(s))$, so that f is rational.

COROLLARY. *A birational isomorphism of the component of the identity* $\mathfrak{G}^0$ *of the Galois group* $\mathfrak{G}$ *onto a group variety can always be extended to a birational isomorphism of* $\mathfrak{G}$ *onto a group variety.*

Proof. In the proposition set $G = \mathfrak{G}$, $G_1 = \mathfrak{G}^0$, and $K =$ the group consisting of the identity automorphism.

Chapter II. Birational isomorphisms between Galois groups and group varieties.

1. Galois groups onto group varieties. In this section we shall prove the following theorem.

THEOREM 1. *The Galois group* $\mathfrak{G}$ *of the strongly normal extension* $\mathcal{G}$ *of the differential field* $\mathcal{F}$ *of characteristic* 0 *with algebraically closed field of constants* $\mathcal{C}$ *is always birationally isomorphic to a group variety defined over* $\mathcal{C}$.

Proof. By the corollary to Proposition 7 of Chapter I, we may assume that $\mathfrak{G}$ is irreducible. To prove the theorem under this added hypothesis, we use a theorem due to Weil ([14], Th. 15, p. 54), and later refined by Rosenlicht ([11], Th. 4, p. 511). To state this theorem we need the following definition due to Weil ([14], p. 51): If W is any irreducible variety, a *normal law of composition* on W is a rational mapping of $W \times W$ into W which, if we denote the law of composition multiplicatively, has the two properties that 1) $\mathcal{C}(s, st) = \mathcal{C}(t, st) = \mathcal{C}(t, s)$ for independent generic points s and t of W, and 2) $(st)u\ s(tu)$ for independent generic points s, t, and u of W. Weil's theorem is: If W is an irreducible variety with a normal law of composition, then there exists a birational correspondence T between W and a group variety G which transforms the normal law of composition on W into the group composition on G (T, G, and the group composition on G not necessarily defined over $\mathcal{C}$). Rosenlicht's contribution is to add: If the points of W which are rational over $\mathcal{C}$ form a dense (Zariski topology) subset of W (which is true in the present case because $\mathcal{C}$ is algebraically closed), then we may take G and T so that T, G, and the group composition on G are defined over $\mathcal{C}$.

Let σ, τ be generic elements of $\mathfrak{G}$ which are independent ([7], pp. 777-778); then $\sigma\tau$ is also a generic element of $\mathfrak{G}$. Let $\gamma_1, \cdots, \gamma_r$ be constants such that $\mathcal{C}(\sigma) = \mathcal{C}(\gamma_1, \cdots, \gamma_r)$. Each γ_i can be expressed as the quotient of two polynomials over $\mathcal{G}$ in a finite number of elements $\sigma\eta_1, \cdots, \sigma\eta_n$, where each $\eta_j \in \mathcal{G}$; fixing two such polynomials, for each generic element ρ of $\mathfrak{G}$ let $c_i(\rho)$ be the element obtained on replacing each $\sigma\eta_j$ by $\rho\eta_j$ in the quotient of the two polynomials. Then $c_i(\sigma) = \gamma_i$, and for each generic element ρ of $\mathfrak{G}$, $\mathcal{C}(\rho) = \mathcal{C}(c_1(\rho), \cdots, c_r(\rho))$. Also, writing $c(\rho) = (c_1(\rho), \cdots, c_r(\rho))$, we see that $c(\sigma)$ is a generic point of a certain variety W in r-dimensional affine space, that $(c(\sigma), c(\tau))$ is a generic point of $W \times W$, and that $c(\sigma\tau)$ is a generic point of W with $\mathcal{C}(c(\sigma\tau)) = \mathcal{C}(\sigma\tau) \subset \mathcal{C}(\sigma, \tau) = \mathcal{C}(c(\sigma), c(\tau))$.

Therefore there exists a rational mapping of $W \times W$ into W which maps $(c(\sigma), c(\tau))$ onto $c(\sigma\tau)$. It is now a simple matter to see that this rational mapping is a normal law of composition on W. Therefore, by the Weil-Rosenlicht result just mentioned, there exists a group variety G, and a birational mapping g' between W and G, such that $g'(c(\sigma\tau)) = g'(c(\sigma))g'(c(\tau))$ whenever σ, τ are independent generic elements of $\mathfrak{G}$. If for each generic element σ of $\mathfrak{G}$ we define $g(\sigma) = g'(c(\sigma))$ then g is a one-to-one mapping of the set of generic elements of $\mathfrak{G}$ onto the set of generic points of G, such that $\mathcal{F}(\sigma) = \mathcal{F}(g(\sigma))$ for every generic element σ of $\mathfrak{G}$, and such that $g(\sigma\tau) = g(\sigma)g(\tau)$ for all pairs σ, τ of independent generic elements of $\mathfrak{G}$; it is easy to see, moreover, that if σ, τ, σ', τ' are generic elements of $\mathfrak{G}$ then (σ', τ') is a specialization of (σ, τ) if and only if $(g(\sigma'), g(\tau'))$ is a specialization of $(g(\sigma), g(\tau))$. Every generic element σ of $\mathfrak{G}$ thus has the property that $g(\sigma\tau)g(\tau)^{-1}$ has the same value for all choices of generic element τ of $\mathfrak{G}$ such that σ, τ are independent. It is now a simple matter to prove that every element σ of $\mathfrak{G}$ has this property. Indeed, if $\sigma \varepsilon \mathfrak{G}$ and if τ, τ' are generic elements of $\mathfrak{G}$ such that σ, τ are independent and σ, τ' are independent, we may find a generic element τ'' of $\mathfrak{G}$ such that σ, τ, τ'' are independent and σ, τ', τ'' are independent; obviously then $\sigma\tau''$, $\tau''^{-1}\tau$ are independent generic points of $\mathfrak{G}$ as are $\tau'', \tau''^{-1}\tau$, so that

$$g(\sigma\tau)g(\tau)^{-1} = g(\sigma\tau'' \cdot \tau''^{-1}\tau)g(\tau'' \cdot \tau''^{-1}\tau)^{-1}$$
$$= g(\sigma\tau'')g(\tau''^{-1}\tau)(g(\tau'')g(\tau''^{-1}\tau))^{-1} = g(\sigma\tau'')g(\tau'')^{-1},$$

and similarly $g(\sigma\tau')g(\tau')^{-1} = g(\sigma\tau'')g(\tau'')^{-1}$, whence $g(\sigma\tau)g(\tau)^{-1} = g(\sigma\tau')g(\tau')^{-1}$.

This being established, we may define a mapping f of $\mathfrak{G}$ into G by the formula $f(\sigma) = g(\sigma\tau)g(\tau)^{-1}$ $(\sigma \varepsilon \mathfrak{G})$, where τ represents any generic element of $\mathfrak{G}$ such that σ, τ are independent. If $\sigma, \sigma' \varepsilon \mathfrak{G}$ we may choose a generic element τ such that each of the three pairs σ, τ and σ', τ and $\sigma, \sigma'\tau$ is independent, so that

$$f(\sigma\sigma') = g(\sigma\sigma'\tau)g(\tau)^{-1} = g(\sigma \cdot \sigma'\tau)g(\sigma'\tau)^{-1} \cdot g(\sigma'\tau)g(\tau)^{-1} = f(\sigma)f(\sigma')$$

and f is a homomorphism of the group $\mathfrak{G}$ into the group G. If ν is any element of the kernel of f and τ is a generic element of $\mathfrak{G}$ such that ν, τ are independent, so that $\nu\tau$ is generic on $\mathfrak{G}$, then $g(\nu\tau) = f(\nu\tau) = f(\tau) = g(\tau)$, whence $\nu\tau = \tau$ and ν is the identity; thus, the kernel of f is trivial. For any $\sigma \varepsilon \mathfrak{G}$ we may find a generic element τ of $\mathfrak{G}$ such that $\mathcal{F}(\mathcal{F}(\sigma), \mathcal{F}(f(\sigma)))$ and $\mathcal{F}(\tau)$ are linearly disjoint over $\mathcal{F}$, whence in particular σ, τ are independent; for such τ we have

$$\mathcal{F}(f(\sigma), g(\tau)) = \mathcal{F}(g(\sigma\tau)g(\tau)^{-1}, g(\tau)) = \mathcal{F}(g(\sigma\tau), g(\tau)) = \mathcal{F}(\sigma\tau, \tau) = \mathcal{F}(\sigma, \tau),$$

whence it easily follows that $\mathcal{K}(f(\sigma)) = \mathcal{K}(\sigma)$. Again, if σ and σ' are any element of $\mathfrak{G}$ and we choose generic elements τ and τ' of $\mathfrak{G}$ such that each pair σ, τ and σ', τ' is independent then ([7], Ch. II, Props. 6, 8, pp. 777, 778) σ' is a specialization of σ if and only if $(\tau'\sigma', \tau')$ is a specialization of $(\tau\sigma, \tau)$; this takes place if and only if $(g(\tau'\sigma'), g(\tau'))$ is a specialization of $(g(\tau\sigma), g(\tau))$, that is, if and only if $(f(\tau')f(\sigma'), f(\tau'))$ is a specialization of $(f(\tau)f(\sigma), f(\tau))$, which in turn occurs if and only if $f(\sigma')$ is a specialization of $f(\sigma)$. To prove that f is a birational isomorphism of $\mathfrak{G}$ onto G it remains to show that G is the image of f. Now, the image contains $f(\tau)$, where τ is a generic element of $\mathfrak{G}$, and $f(\tau)$ is a generic element of G; therefore the smallest subvariety of G conaining the image is G itself. It follows (with the help of [7], Ch. II, Prop. 9, p. 779) that the complement of the image in G is contained in a proper subvariety of G, so that (Ch. I, Prop. 3) the image of f is G.

2. Irreducible group variety onto Galois group.

We now prove the following partial converse to Theorem 1.

THEOREM 2. *Every irreducible group variety defined over the algebraically closed field $\mathcal{K}$ of characteristic 0 is birationally isomorphic to the Galois group of a strongly normal extension of a differential field with field of constants $\mathcal{K}$.*

Proof. Let G be an irreducible group variety defined over $\mathcal{K}$, and set $r = \dim G$. The field $\mathcal{H}_0^*$ of all rational functions on G (not necessarily defined over $\mathcal{K}$) is a finitely generated field extension of transcendence degree r of its subfield of all constant functions on G, which subfield we identify with $\mathcal{K}^*$ by identifying each constant function with its value. The subfield $\mathcal{H}_0$ of $\mathcal{H}_0^*$ consisting of all rational functions on G for which $\mathcal{K}$ is a field of definition, is a finitely generated field extension of transcendence degree r of $\mathcal{K}$. $\mathcal{H}_0$ and $\mathcal{K}^*$ are linearly disjoint over $\mathcal{K}$ (so that their intersection is $\mathcal{K}$), and their compositum is $\mathcal{H}_0^*$.

For each point $s \in G$ there exists a unique automorphism ρ_s of the field $\mathcal{H}_0^*$ with the following property: if $f \in \mathcal{H}_0^*$ and $t \in G$ then $\rho_s f$ is defined at t if and only if f is defined at ts and, when this is the case, $(\rho_s f)(t) = f(ts)$. The computation

$$((\rho_s\rho_{s'})f)(t) = (\rho_s(\rho_{s'}f))(t) = (\rho_{s'}f)(ts) = f(tss') = (\rho_{ss'}f)(t)$$

shows that

$$\rho_{ss'} = \rho_s\rho_{s'} \quad (s, s' \in G). \tag{1}$$

Also, if $\rho_s f = f$ for every $f \in \mathcal{H}_0^*$, that is if $f(ts) = f(t)$ for all $f \in \mathcal{H}_0^*$ and a generic point t of G such that t, s are independent (over $\mathcal{K}$), then obviously s is the unity element of G, that is

(2) "$\rho_s =$ identity automorphism" implies "$s =$ unity element of G."

If $s \in G$, and if t denotes any point of G (in particular, any generic point of G such that s, t are independent over $\mathcal{K}$), then $\mathcal{K}(t, ts) = \mathcal{K}(t, s) = \mathcal{K}(ts, s)$. It follows that

$$\mathcal{U}_0(\rho_s \mathcal{U}_0) = \mathcal{U}_0(\mathcal{K}(s)) = (\rho_s \mathcal{U}_0)(\mathcal{K}(s)). \tag{3}$$

Also if $s' \in G$ and we take t generic with s, t independent and s', t independent then s' is a specialization of s over $\mathcal{K}$ if and only if s' is a specialization of s over $\mathcal{K}(t)$, which occurs if and only if $(f(ts'))_{f \in \mathcal{H}_0}$ is a specialization of of $(f(ts))_{f \in \mathcal{H}_0}$ over $\mathcal{K}(t)$, that is, $((\rho_{s'} f)(t))_{f \in \mathcal{H}_0}$ is a specialization of $((\rho_s f)(t))_{f \in \mathcal{H}_0}$ over $\mathcal{K}(t)$; it follows that

$$\begin{cases} s' \text{ is a specialization of } s \text{ over } \mathcal{K} \text{ if and only if} \\ (\rho_{s'} f)_{f \in \mathcal{H}_0} \text{ is a specialization of } (\rho_s f)_{f \in \mathcal{H}_0} \text{ over } \mathcal{U}_0. \end{cases} \tag{4}$$

The set of all derivations of $\mathcal{U}_0^*$ over $\mathcal{K}^*$ is a vector space over $\mathcal{U}_0^*$ of dimension r. It is known (for example, see Nakano [9]) that this space has a base $D_1, \cdots, D_r$ composed of derivations each of which maps $\mathcal{U}_0$ into itself and is "right-invariant" in the sense that

$$\rho_s \circ D_i = D_i \circ \rho_s \qquad (s \in G, 1 \leqq i \leqq r). \tag{5}$$

An element of $\mathcal{U}_0^*$ annihilated by $D_1, \cdots, D_r$ is annihilated by every derivation of $\mathcal{U}_0^*$ over $\mathcal{K}^*$, and therefore belongs to $\mathcal{K}^*$.

Now let $(u_{ij})_{1 \leqq i \leqq r, 0 \leqq j < \infty}$ be a family of indeterminates, and let

$$\mathcal{E} = \mathcal{K}((u_{ij})_{1 \leqq i \leqq r, 0 \leqq j < \infty}), \quad \mathcal{U} = \mathcal{U}_0((u_{ij})_{1 \leqq i \leqq r, 0 \leqq j < \infty});$$

the fields $\mathcal{K}^*((u_{ij})_{1 \leqq i \leqq r, 0 \leqq j < \infty})$ and $\mathcal{U}_0^*((u_{ij})_{1 \leqq i \leqq r, 0 \leqq j < \infty})$ may then be denoted by $\mathcal{E}(\mathcal{K}^*)$ and $\mathcal{U}(\mathcal{K}^*)$, respectively. For each $s \in G$ the automorphism ρ_s of $\mathcal{U}_0^*$ over $\mathcal{K}^*$, can be extended to a unique automorphism (which we also denote by ρ_s) of $\mathcal{U}(\mathcal{K}^*)$ over $\mathcal{E}(\mathcal{K}^*)$. For each i the invariant derivation D_i of $\mathcal{U}_0^*$ over $\mathcal{K}^*$ can be extended to a unique derivation (which we also denote by D_i) of $\mathcal{U}(\mathcal{K}^*)$ over $\mathcal{K}^*$ with the property that $D_i u_{ij} = u_{i0}^{-1} u_{i,j+1}$ and $D_i u_{i'j} = 0$ whenever $i' \neq i$ $(0 \leqq j < \infty)$. For the extended automorphisms and derivations (1), (2), (5) continue to hold, and (3), (4) yield

$$\mathcal{U}(\rho_s \mathcal{U}) = \mathcal{U}(\mathcal{K}(s)) = (\rho_s \mathcal{U})(\mathcal{K}(s)) \qquad (s \in G) \tag{6}$$

$$\begin{cases} s' \text{ is a specialization of } s \text{ over } \mathcal{K} \text{ if and only if} \\ (\rho_{s'} \alpha)_{\alpha \in \mathcal{H}} \text{ is a specialization of } (\rho_s \alpha)_{\alpha \in \mathcal{H}} \text{ over } \mathcal{U}. \end{cases} \tag{7}$$

Furthermore, if α is any element of $\mathcal{H}(\mathcal{C}^*)$, then α may be written as a rational fraction with coefficients in $\mathcal{H}_0^*$ in the indeterminate u_{ij}, and if α contains u_{kl} effectively but for every $l' > l$ fails to contain $u_{kl'}$ effective then it is clear that $u_{k,l+1}$ is effectively present in $D_k\alpha$, so that $D_k\alpha \neq 0$; thus the only elements of $\mathcal{H}(\mathcal{C}^*)$ annihilated by every D_i are the elements of $\mathcal{C}^*$:

$$D_i\alpha = 0\,(1 \leqq i \leqq r) \text{ if and only if } \alpha \in \mathcal{C}^*. \tag{8}$$

We now make $\mathcal{H}(\mathcal{C}^*)$ into a differential field (not necessarily contained in the universal extension $\mathcal{F}^*$ of $\mathcal{G}$ introduced in Section 1 of Chapter I) with derivation operators $\delta_1, \cdots, \delta_m$ by defining

$$\delta_1\alpha = \sum u_{i0} D_i\alpha, \qquad \delta_k\alpha = 0\,(2 \leqq k \leqq m) \tag{9}$$

for every $\alpha \in \mathcal{H}(\mathcal{C}^*)$. It is clear that $\mathcal{C}$, $\mathcal{E}$, $\mathcal{H}$, $\mathcal{C}^*$ all are differential subfields of $\mathcal{H}(\mathcal{C}^*)$, so that $\mathcal{H}(\mathcal{C}^*)$ may now be written as $\mathcal{H}\langle\mathcal{C}^*\rangle$; furthermore, if we denote u_{i0} by u_i then $u_{ij} = \delta_1{}^j u$ $(1 \leqq i \leqq r, 0 \leqq j < \infty)$, so that we may write $\mathcal{E} = \mathcal{C}\langle u_1, \cdots, u_r\rangle$ and $\mathcal{H} = \mathcal{E}\langle\mathcal{H}_0\rangle$.

If $\alpha \in \mathcal{H}\langle\mathcal{C}^*\rangle$ contains u_{kl} effectively but for every $l' > l$ fails to contain $u_{kl'}$ effectively, then $\delta_1\alpha = \sum u_{i0} D_i\alpha$ contains $u_{k,l+1}$ effectively, so that $\delta_1\alpha \neq 0$. Therefore if $\delta_i\alpha = 0$ then $\alpha \in \mathcal{H}_0^*$, and from the equation $\sum u_{i0} D_i\alpha = 0$ and the fact that $u_{10}, \cdots, u_{r0}$ are linearly independent over $\mathcal{H}_0^*$ we conclude that $D_i\alpha = 0$ $(1 \leqq i \leqq r)$. It follows from (8) that the field of constants of the differential field $\mathcal{H}\langle\mathcal{C}^*\rangle$ is $\mathcal{C}^*$; $\mathcal{C}^*$ is therefore also the field of constants of $\mathcal{E}\langle\mathcal{C}^*\rangle$, and $\mathcal{C}$ is the field of constants of $\mathcal{H}$ and of $\mathcal{E}$.

Before proceeding farther, we rid ourselves of the awkward circumstance that $\mathcal{H}\langle\mathcal{C}^*\rangle$ is not in general a differential subfield of $\mathcal{F}^*$. In the first place, $\mathcal{H} = \mathcal{E}\langle\mathcal{H}_0\rangle = \mathcal{C}\langle u_1, \cdots, u_r, \mathcal{H}_0\rangle$, so that $\mathcal{H}$ is a finitely generated differential field extension of $\mathcal{C}$; since $\mathcal{F}^*$ is a universal extension of $\mathcal{C}$ (see [7], Ch. I, § 5, p. 768), there exists an isomorphism over $\mathcal{C}$ of $\mathcal{H}$ onto a differential subfield $\mathcal{H}'$ of $\mathcal{F}^*$. Secondly, a family of elements of $\mathcal{H}$ linearly independent over $\mathcal{C}$ is also linearly independent over $\mathcal{C}^*$ (see [7], Ch. I, § 4, Prop. 3 and remark following, pp. 766-767), so that $\mathcal{H}$ and $\mathcal{C}^*$ are linearly disjoint over $\mathcal{C}$; it follows that the isomorphism over $\mathcal{C}$ of $\mathcal{H}$ onto $\mathcal{H}'$ can be extended to a unique isomorphism over $\mathcal{C}^*$ of $\mathcal{H}\langle\mathcal{C}^*\rangle$ onto $\mathcal{H}'\langle\mathcal{C}^*\rangle$. By virtue of this isomorphism we identify $\mathcal{H}\langle\mathcal{C}^*\rangle$ with $\mathcal{H}'\langle\mathcal{C}^*\rangle$, so that henceforth all our differential fields are differential subfields of $\mathcal{F}^*$.

By (5) and (9) we know that $\rho_s\delta_i\alpha = \delta_i\rho_s\alpha$ $(s \in G, 1 \leqq i \leqq m, \alpha \in \mathcal{H}\langle\mathcal{C}^*\rangle)$, so that each ρ_s is an automorphism of $\mathcal{H}\langle\mathcal{C}^*\rangle$ as a differential field (not merely as a field). This proves ([7], Ch. II, § 3, Prop. 4, p. 775) that each ρ_s, identified with its restriction to $\mathcal{H}$, is a strong isomorphism of $\mathcal{H}$ over $\mathcal{E}$.

It follows from (6) that if t is a generic point of G then the transcendence degree of $\mathfrak{M}\langle\rho_t\mathfrak{M}\rangle$ over $\mathfrak{M}$ equals r, which in turn equals the transcendence degree of $\mathfrak{M}$ over $\mathcal{E}$. Consequently ([7], Ch. II, § 2, Prop. 2, p, 773) ρ_t is an isolated isomorphism of $\mathfrak{M}$ over $\mathcal{E}$. Since $\mathcal{E}$ is obviously relatively algebraically closed in $\mathfrak{M}$, it follows ([7], Ch. II, § 2, Prop. 3, p. 774) that every isomorphism of $\mathfrak{M}$ over $\mathcal{E}$ is a specialization of ρ_t, and therefore ([7], Ch. II, § 4, Prop. 5, p. 776) is strong. Thus $\mathfrak{M}$ is a strongly normal extension of $\mathcal{E}$; we denote its Galois group by $\mathfrak{H}$.

By (1), the mapping $f: s \rightarrow \rho_s$ $(s \in G)$ is a homomorphism of G into $\mathfrak{H}$. By (2) the kernel of f is trivial. By (6) and [7], Ch. I, § 4, Cor. 5 to Prop. 3, p. 768,

$$\mathcal{K}(\rho_s) = \mathcal{K}(s) \qquad (s \in G),$$

and by (7), $\rho_{s'}$ is a specialization of ρ_s if and only if s' is a specialization of s (over $\mathcal{K}$). Therefore f is a rational homomorphism of G into $\mathfrak{H}$. Since the Galois group $\mathfrak{H}$ is by Theorem 1 birationally isomorphic with a group variety, it follows from the corollary to Proposition 5 in Chapter I that the image of f is an algebraic subgroup of $\mathfrak{H}$. Since the image contains a generic point $f(t)$ of $\mathfrak{H}$, it follows that the image of f is $\mathfrak{H}$ itself. Again making use of the fact that $\mathfrak{H}$ is birationally isomorphic with a group variety, we conclude from the corollary to Proposition 6 in Chapter I that f is a birational isomorphism of G onto $\mathfrak{H}$.

3. Applications. It follows from Theorem 1 that many propositions about group varieties can be generalized by permitting each group variety in the statement of a given proposition to be any algebraic group (that is, either a Galois group or a group variety). For example (as we have already had occasion to observe in Section 2):

PROPOSITION 1. *Proposition 5 of Chapter I and its corollary, and Proposition 6 of Chapter I and its corollary can be generalized to arbitrary algebraic groups.*

Another example:

PROPOSITION 2. *If* $\sigma_1, \cdots, \sigma_n \in \mathfrak{G}$ *and if* $(\sigma_1', \cdots, \sigma_n')$ *is a specialization of* $(\sigma_1, \cdots, \sigma_n)$, *then* $(W_i(\sigma'_1, \cdots, \sigma_n'))_{i \in I}$ *is a specialization of* $(W_i(\sigma_1, \cdots, \sigma_n))_{i \in I}$, *where* $(W_i(X_1, \cdots, X_n))_{i \in I}$ *is any family of words in* n *indeterminates* $X_1, \cdots, X_n$.

This is an immediated consequence of the fact that the same proposition holds for group varieties.

On the other hand some propositions on group varieties can not be indiscriminately generalized in this way because such a generalization may depend on Theorem 2, which we have proved only for *irreducible* group varieties. For example, we have not the right to assert the generalization of Proposition 7 of Chapter I, or of its corollary.

It is perhaps of interest to observe that not only can we prove results on Galois groups by appealing to previously proved analogous results on group varieties, but also we can work in the opposite direction. For example, we can prove that, roughly speaking, a factor group of a group variety is itself a group variety. More precisely:

PROPOSITION 3. *If G is a group variety and K is a normal algebraic subgroup of G then, there exists a rational homomorphism of G with kernel K and image a group variety.*

This result is contained (except for the rather trivial extension here to group varieties which are not necessarily irreducible) in much stronger results of Nakano [10]; the novelty here lies in the fact that the result on group varieties is obtained as a consequence of theorems on algebraic differential equations. The proof runs as follows. By Proposition 1 of Chapter I, the result holds for a group variety G and kernel K provided it holds for the group variety G^0 (= the component of the unity of G) and kernel $G^0 \cap K$, that is, we may assume that G is irreducible. Making this assumption, we know from Theorem 2 that G is birationally isomorphic with a Galois group, say f is a birational isomorphism of G onto $\mathfrak{G}$; $f(K)$ is a *normal* algebraic subgroup of $\mathfrak{G}$ and therefore is the Galois group $\mathfrak{G}(\mathfrak{M})$ of $\mathfrak{G}$ over a differential field $\mathfrak{M}$ intermediate to $\mathfrak{F}$ and $\mathfrak{G}$ and *strongly normal* over $\mathfrak{F}$. Therefore (see Example 1 of Chapter I, § 4) the mapping r, which to each $\sigma \varepsilon \mathfrak{G}$ corresponds the restriction of σ to $\mathfrak{M}$, is a rational homomorphism of $\mathfrak{G}$ with kernel $\mathfrak{G}(\mathfrak{M})$ and image the Galois group $\mathfrak{H}$ of $\mathfrak{M}$ over $\mathfrak{F}$. By Theorem 1, there exists a birational isomorhpism h of $\mathfrak{H}$ onto a group variety H. It is now immediate that the composite mapping $h \circ r \circ f$ is a rational homomorphism of G with kernel K and image H.

Similarly, we can prove the following result.

PROPOSITION 4. *If a rational homomorphism of an algebraic group G has kernel K and image H, then* $\dim G = \dim K + \dim H$.

This too is contained in results of Nakano [10]. Our proof uses the fact that the dimension of the Galois group of a strongly normal differential field extension equals the transcendence degree of the extension. We omit the details, which are rather obvious.

Finally, we generalize to arbitrary connected group varieties (but only in the case of field characteristic 0, of course) a theorem on algebraic matric groups (Chevalley [3], Ch. II, § 4, Th. 1, part 2), p. 84, or Chevalley and Kolchin [4], § 1, Th. 2).

Proposition 5. *The notation being as in § 2, if W is a subset of an irreducible group variety G then a necessary and sufficient condition for W to be an algebraic subgroup of G is that there exist a finite subset S, of the field $\mathcal{H}_0$ of rational functions on G which are defined over $\mathcal{K}$, such that W is the set of all points $s \varepsilon G$ for which $\rho_s\alpha = \alpha$ for every $\alpha \varepsilon S$.*

Indeed, we saw in the proof of Theorem 2 that $\mathcal{H} = \mathcal{E}\langle\mathcal{H}_0\rangle$ is a strongly normal extension of $\mathcal{E} = \mathcal{K}\langle u_1, \cdots, u_r\rangle$, and $s \rightarrow \rho_s$ $(s \varepsilon G)$ is a birational isomorphism f of G onto the Galois group $\mathfrak{H}$ of $\mathcal{H}$ over $\mathcal{E}$. W is an algebraic subgroup of G if and only if $f(W)$ is an algebraic subgroup of $\mathfrak{H}$, that is, is the Galois group of $\mathcal{H}$ over a differential field $\mathcal{E}_1$ intermediate to $\mathcal{E}$ and $\mathcal{H}$; now every intermediate differential field is a finitely generated differential field extension, and it is easy to see that $\mathcal{E}_1 = \mathcal{E}\langle\alpha_1, \cdots, \alpha_p\rangle$ for some finite set S of elements $\alpha_1, \cdots, \alpha_p$ which belong to $\mathcal{H}_0$. Therefore $f(W)$ is the Galois group of $\mathcal{H}$ over $\mathcal{E}_1$ if and only if W is the set of all elements $s \varepsilon G$ such that $\rho_s\alpha_i = \alpha_i$ $(1 \leqq i \leqq p)$.

Chapter III. One-dimensional cohomology.

1. **Definition of $H^1(\mathfrak{G}, G)$.** The universal extension $\mathcal{F}^*$ of $\mathcal{G}$ is, like its field of constants $\mathcal{K}^*$, an algebraically closed field extension of $\mathcal{K}$ of infinite transcendence degree and therefore may be used, like $\mathcal{K}^*$, as a universal domain for algebraic geometry. Furthermore, since $\mathcal{K}^* \subset \mathcal{F}^*$, every variety V in the algebraic geometry based on $\mathcal{K}^*$ as universal domain determines in an obvious way a variety $V^\dagger$ in the algebraic geometry based on $\mathcal{F}^*$ as universal domain, such that a point of $V^\dagger$ belongs to V if and only if the point is rational over $\mathcal{K}^*$. It is obvious that if V is a group variety in the algebraic geometry based on $\mathcal{K}^*$ then $V^\dagger$ is a group variety in the algebraic geometry based on $\mathcal{F}^*$.

Now, a mapping f of the Galois group $\mathfrak{G}$ of $\mathcal{G}$ over $\mathcal{F}$ into a group variety G may also be regarded as a mapping of $\mathfrak{G}$ into $G^\dagger$; since $\mathcal{K}(f(\sigma))$ consists of constants for any $\sigma \varepsilon \mathfrak{G}$, the condition "$\mathcal{K}(f(\sigma)) \subset \mathcal{K}(\sigma)$" is equivalent to "$\mathcal{G}(f(\sigma)) \subset \mathcal{G}(\sigma\mathcal{G})$," for any $\sigma, \sigma' \varepsilon \mathfrak{G}$ the condition "$f(\sigma')$ is a specialization of $f(\sigma)$ over $\mathcal{K}$" is equivalent to the condition "$f(\sigma')$ is a specialization of $f(\sigma)$ over $\mathcal{G}$" and the condition "$f(\sigma\sigma') = f(\sigma)f(\sigma')$" may, if we wish, be written in the form "$f(\sigma\sigma') = f(\sigma) \cdot \sigma f(\sigma')$." Therefore the following definition generalizes the notion of rational homomorphism of $\mathfrak{G}$ into G:

By a *rational crossed homomorphism*, or a *cocycle of dimension* 1, or a 1-*cocycle*, of the Galois group $\mathfrak{G}$ into the group variety G we mean any mapping f of $\mathfrak{G}$ *into* $G^\dagger$ such that $\mathcal{G}(f(\sigma)) \subset \mathcal{G}(\sigma\mathcal{G})$ for all $\sigma \varepsilon \mathfrak{G}$, such that "$\sigma'$ is a specialization of σ" implies "$f(\sigma')$ is a specialization of $f(\sigma)$ over $\mathcal{G}$" for all $\sigma, \sigma' \varepsilon \mathfrak{G}$, and such that $f(\sigma\sigma') = f(\sigma)\cdot\sigma f(\sigma')$ for all $\sigma, \sigma' \varepsilon \mathfrak{G}$.

As remarked above, every rational homomorphism of $\mathfrak{G}$ into G is a rational crossed homomorphism of $\mathfrak{G}$ into G. Also, if α is any point of $G^\dagger$ *which is rational over* $\mathcal{G}$, then it is evident that the mapping $\sigma \rightarrow \alpha^{-1}\sigma\alpha$ $(\sigma \varepsilon \mathfrak{G})$ is a rational crossed homomorphism of $\mathfrak{G}$ into G; such a rational crossed homomorphism is said to be *trivial*, or to be a *coboundary of dimension* 1 or a 1-*coboundary* of $\mathfrak{G}$ into G.

We denote the set of all 1-cocyles of $\mathfrak{G}$ into G by $Z^1(\mathfrak{G}, G)$, and the set of all 1-coboundaries by $B^1(\mathfrak{G}, G)$, so that $B^1(\mathfrak{G}, G) \subset Z^1(\mathfrak{G}, G)$. The mapping which to each element of $\mathfrak{G}$ corresponds the unity element 1 of G is obviously a 1-coboundary (because $1^{-1}\sigma 1 = 1$ for every $\sigma \varepsilon \mathfrak{G}$); we denote it by 1. In general, as we have made no assumption of commutativity for G, $Z^1(\mathfrak{G}, G)$ and $B^1(\mathfrak{G}, G)$ are not groups under the law of composition $(f, g) \rightarrow fg$, where $(fg)(\sigma) = f(\sigma)g(\sigma)$, because fg need not be an element of $Z^1(\mathfrak{G}, G)$.

If $f \varepsilon Z^1(\mathfrak{G}, G)$, and if $\alpha \varepsilon G^\dagger$ is rational over $\mathcal{G}$, then the mapping g of $\mathfrak{G}$ defined by $g(\sigma) = \alpha^{-1}f(\sigma)\sigma\alpha$ is obviously an element of $Z^1(\mathfrak{G}, G)$, which we call *cohomologous* to f. The relation "g is cohomologous to f" is an equivalence on $Z^1(\mathfrak{G}, G)$. An equivalence class with respect to this equivalence we call a *cohomology class of dimension* 1 or a 1-*cohomology class* of $\mathfrak{G}$ into G; we denote the set of all these by $H^1(\mathfrak{G}, G)$. $B^1(\mathfrak{G}, G)$ itself is a 1-cohomology class; when we think of this class as an element of $H^1(\mathfrak{G}, G)$ we frequently permit ourselves to denote it by 1, and we also permit ourselves to write $H^1(\mathfrak{G}, G) = 1$ when we wish to assert that $B^1(\mathfrak{G}, G)$ is the only 1-cohomology class of $\mathfrak{G}$ into G.

Of course, if G is commutative then $Z^1(\mathfrak{G}, G)$ and $B^1(\mathfrak{G}, G)$ are commutative groups, and $H^1(\mathfrak{G}, G)$ is the factor group $Z^1(\mathfrak{G}, G)/B^1(\mathfrak{G}, G)$; we then call it the *first cohomology group* of $\mathfrak{G}$ into G.

2. Determination of $H^1(\mathfrak{G}, G)$ in some general cases. The computation of $H^1(\mathfrak{G}, G)$ in various cases below is facilitated by the observation that two 1-cocycles f and g of $\mathfrak{G}$ into G are equal provided $f(\sigma) = g(\sigma)$ for all $\sigma \varepsilon \mathfrak{G}$ which are automorphisms of $\mathcal{G}$ over $\mathcal{F}$); this follows from [7], Ch. II, Cor. 2 to Prop. 9, p. 783.

Our results on $H^1(\mathfrak{G}, G)$ are summarized in the following theorem. We recall that $\mathfrak{G}$ is the Galois group of the strongly normal extension $\mathcal{G}$

of the differential field $\mathcal{F}$ of characteristic 0 with algebraically closed field of constants $\mathcal{C}$. G denotes any group variety defined over $\mathcal{C}$. We denote the group of all invertible square matrices of degree n with constant coordinates by $GL(n)$; $GL(n)$ is an irreducible group variety of dimension n^2.

THEOREM 1. 1) $H^1(\mathfrak{G}, GL(n)) = 1$.

2) *If* $\mathcal{F}$ *is algebraically closed then* $H^1(\mathfrak{G}, G) = 1$.

3) *If* G *is commutative then every element of the group* $H^1(\mathfrak{G}, G)$ *is of finite order.*

Proof of 1). Let $f \varepsilon Z^1(\mathfrak{G}, GL(n))$, and let τ be a generic element of the component of the identity $\mathfrak{G}^0$ of $\mathfrak{G}$. Since $\mathcal{G}(f(\tau)) \subset \mathcal{G}(\tau\mathcal{G})$, there must exist a finite number of elements $\tau\xi_1, \cdots, \tau\xi_p \varepsilon \tau\mathcal{G}$ which are algebraically independent over $\mathcal{G}$, and a finite number of elements $\tau\eta_1, \cdots, \tau\eta_q \varepsilon \tau\mathcal{G}$ which are linearly independent over $\mathcal{G}(\tau\xi_1, \cdots, \tau\xi_p)$, such that

$$\mathcal{G}(f(\tau)) \subset \sum_{j=1}^{q} \mathcal{G}(\tau\xi_1, \cdots, \tau\xi_p) \cdot \tau\eta_j.$$

Therefore there exist matrices $A_j = A_j(X_1, \cdots, X_p)$, $1 \leqq j \leqq q$, square and of degree n, with coordinates which are polynomials over $\mathcal{G}$ in indeterminates $X_1, \cdots, X_p$, and there exists a nonzero polynomial

$$B = B(X_1, \cdots, X_p) \varepsilon \mathcal{G}[X_1, \cdots, X_p],$$

such that

$$f(\tau) = \sum_{j=1}^{q} (\tau\eta_j) B(\tau\xi_1, \cdots, \tau\xi_p)^{-1} A_j(\tau\xi_1, \cdots, \tau\xi_p). \tag{1}$$

This equation continues to hold, of course, if τ is replaced by any generic element of $\mathfrak{G}^0$. We assume without loss of generality that B does not have a factor of degree > 0 which divides every coordinate of every A_j, and we assume, too, that one of the coefficients in B is 1. For any $\sigma \varepsilon \mathfrak{G}^0$ which is an *auto*morphism of $\mathcal{G}$, $\sigma^{-1}\tau$ is a generic element of $\mathfrak{G}^0$. On the other hand, $f(\tau) = f(\sigma \cdot \sigma^{-1}\tau) = f(\sigma) \cdot \sigma f(\sigma^{-1}\tau)$. Therefore by (1)

$$\begin{aligned} &\sum_{j=1}^{q} (\tau\eta_j) B(\tau\xi_1, \cdots, \tau\xi_p)^{-1} A_j(\tau\xi_1, \cdots, \tau\xi_p) \\ &\qquad = f(\sigma) \cdot \sum_{j=1}^{q} (\tau\eta_j) B_\sigma(\tau\xi_1, \cdots, \tau\xi_p)^{-1} A_{j\sigma}(\tau\xi_1, \cdots, \tau\xi_p), \end{aligned}$$

where B_σ is obtained from B by replacing each coefficient by its image under σ, and similarly for $A_{j\sigma}$. Because $\tau\eta_1, \cdots, \tau\eta_q$ are linearly independent over $\mathcal{G}(\tau\xi_1, \cdots, \tau\xi_p)$, it follows that

$$\begin{aligned} &B(\tau\xi_1, \cdots, \tau\xi_p)^{-1} A_j(\tau\xi_1, \cdots, \tau\xi_p) \\ &\qquad = f(\sigma) B_\sigma(\tau\xi_1, \cdots, \tau\xi_p)^{-1} A_{j\sigma}(\tau\xi_1, \cdots, \tau\xi_p), \qquad 1 \leqq j \leqq p. \end{aligned}$$

Because $\tau\xi_1, \cdots, \tau\xi_p$ are algebraically independent over $\mathcal{G}$, it follows that

$$B_\sigma A_j = f(\sigma) \cdot BA_{j\sigma}, \qquad 1 \leqq j \leqq p.$$

Because no factor of B of degree > 0 divides every coordinate of every A_j, it follows that B_σ is a multiple of B by an element of $\mathcal{G}$, and because a coefficient in B is 1 we conclude that $B_\sigma = B$, so that

$$(2) \qquad A_j = f(\sigma) A_{j\sigma}, \qquad 1 \leqq j \leqq p.$$

Now, it is a consequence of (1) that if $Y_1, \cdots, Y_q$ are additional indeterminates then

$$\det\left(\sum_{j=1}^{q} Y_j A_j(X_1, \cdots, X_p)\right) \neq 0.$$

Therefore there exist element $a_1, \cdots, a_p, b_1, \cdots, b_q \varepsilon \mathcal{F}$ such that if we form the matrix $\alpha = \sum b_j A_j(a_1, \cdots, a_p)$ then $\det \alpha \neq 0$; obviously the coordinates of α are elements of $\mathcal{G}$. However,

$$\sigma A_j(a_1, \cdots, a_p) = A_{j\sigma}(\sigma a_1, \cdots, \sigma a_p) = A_{j\sigma}(a_1, \cdots, a_p).$$

Therefore by (2) we have $\alpha = f(\sigma)\sigma\alpha$. Thus, if we let g be the 1-cocycle cohomologous to f defined by the equation $g(\sigma') = \alpha^{-1} f(\sigma')\sigma'\alpha$ for all $\sigma' \varepsilon \mathfrak{G}$, then $g(\sigma) = 1$ for every $\sigma \varepsilon \mathfrak{G}^0$.

If σ' and σ'' belong to the same coset of $\mathfrak{G}$ relative to $\mathfrak{G}^0$, then $g(\sigma'') = g(\sigma' \cdot \sigma'^{-1}\sigma'') = g(\sigma')\sigma' g(\sigma'^{-1}\sigma'') = g(\sigma')$, so that g is constant on each coset. Since every coset contains an automorphism of $\mathcal{G}$, $g(\sigma')$ is always rational over $\mathcal{G}$; but, for any $\sigma \varepsilon \mathfrak{G}^0$, we have $g(\sigma') = g(\sigma\sigma') = g(\sigma)\sigma g(\sigma') = \sigma g(\sigma')$, so that $g(\sigma')$, is always rational over the relative algebraic closure $\mathcal{F}^0$ of $\mathcal{F}$ in $\mathcal{G}$. Of course, if $\sigma_1, \cdots, \sigma_h$ form a complete set of representatives of the cosets of $\mathfrak{G}$ relative to $\mathfrak{G}^0$, then the restrictions of $\sigma_1, \cdots, \sigma_h$ to $\mathcal{F}^0$ are precisely the elements of the Galois group of $\mathcal{F}^0$ over $\mathcal{F}$. Now, if $T_1, \cdots, T_h$ are indeterminates, then $\det(\sum T_i g(\sigma_i))$ is a polynomial in $\mathcal{F}^0[T_1, \cdots, T_h]$, and is different from 0 (because, for example, at $(1, 0, \cdots, 0)$ it has the value $\det(g(\sigma_1)))$. Therefore (Bourbaki [1], Ch. V, § 10, Th. 4, p. 157) there exists an element $a \varepsilon \mathcal{F}^0$ such that $\det(\sum(\sigma_i a) g(\sigma_i)) \neq 0$. Writing $\beta = \sum(\sigma_i a) g(\sigma_i)$, we have for any $\sigma' \varepsilon \mathfrak{G}$,

$$\sigma'\beta = \sum(\sigma'\sigma_i a)\sigma' g(\sigma_i) = \sum(\sigma'\sigma_i a) g(\sigma')^{-1} g(\sigma'\sigma_i) = g(\sigma')^{-1}\beta.$$

Letting $\gamma = \beta^{-1}\alpha^{-1}$ we therefore obtain $\gamma^{-1}\sigma'\gamma = \alpha\beta\sigma'\beta^{-1} \cdot \sigma'\alpha^{-1} = \alpha g(\sigma')\sigma'\alpha^{-1} = f(\sigma')$ for every $\sigma' \varepsilon \mathfrak{G}$, so that $f \varepsilon B^1(\mathfrak{G}, GL(n))$.

Proof of 2). Since $\mathcal{F}$ is algebraically closed $\mathfrak{G}$ is irreducible. Let $f \varepsilon Z^1(\mathfrak{G}, G)$, and let τ be a generic element of $\mathfrak{G}$. Letting $\eta_1, \cdots, \eta_n$ be elements of $\mathcal{G}$ such that $\mathcal{G} = \mathcal{F}(\eta_1, \cdots, \eta_n)$, we see that

$$\mathcal{G}(\tau\mathcal{G}) = \mathcal{F}(\eta_1, \cdots, \eta_n, \tau\eta_1, \cdots, \tau\eta_n);$$

also, $(\eta_1, \cdots, \eta_n)$ and $(\tau\eta_1, \cdots, \tau\eta_n)$ are independent over $\mathfrak{F}$ (because the transcendence degree of $\mathfrak{F}(\eta_1, \cdots, \eta_n, \tau\eta_1, \cdots, \tau\eta_n)$ over $\mathfrak{F}$ equals that of $\mathfrak{F}(\eta_1, \cdots, \eta_n)$ over $\mathfrak{F}$ plus that of $\mathfrak{F}(\tau\eta_1, \cdots, \tau\eta_n)$ over $\mathfrak{F}$).

Let $f_i(\tau)$ denote a representative in affine space of the abstract point $f(\tau)$. Because $f(\tau)$ is rational over $\mathfrak{G}(\tau\mathfrak{G})$ there exist rational fractions $R_{i1}, \cdots, R_{id_i}$ in $2n$ indeterminates, with coefficients in $\mathfrak{F}$ and denominators different from 0 at $(\eta_1, \cdots, \eta_n, \tau\eta_1, \cdots, \tau\eta_n)$, such that

$$f_i(\tau) = (R_{i1}(\eta_1, \cdots, \eta_n, \tau\eta_1, \cdots, \tau\eta_n), \cdots, R_{id_i}(\eta_1, \cdots, \eta_n, \tau\eta_1, \cdots, \tau\eta_n))$$

For any $\sigma \in \mathfrak{S}$ which is an *auto*morphism of $\mathfrak{G}$, the corresponding representative of $\sigma f(\sigma^{-1}\tau)$ is obviously given by

$$\sigma f_i(\sigma^{-1}\tau) = (R_{i1}(\sigma\eta_1, \cdots, \sigma\eta_n, \tau\eta_1, \cdots, \tau\eta_n), \cdots, R_{id_i}(\sigma\eta_1, \cdots, \sigma\eta_n, \tau\eta_1, \cdots, \tau\eta_n)).$$

Since $\mathfrak{F}$ is algebraically closed, there exists an affine point $(b_1, \cdots, b_n)$ which is a specialization in the sense of algebraic geometry of $(\tau\eta_1, \cdots, \tau\eta_n)$ over $\mathfrak{F}$, with each $b_i \in \mathfrak{F}$ and with the denominator of each R_{ij} different from 0 at $(\eta_1, \cdots, \eta_n, b_1, \cdots, b_n)$; application of the automorphism σ to $R_{ij}(\eta_1, \cdots, \eta_n, b_1, \cdots, b_n)$ shows that the denominator of each R_{ij} is different from 0 at $(\sigma\eta_1, \cdots, \sigma\eta_n, b_1, \cdots, b_n)$. Since $(\eta_1, \cdots, \eta_n)$ and $(\tau\eta_1, \cdots, \tau\eta_n)$ are independent over $\mathfrak{F}$, $(\eta_1, \cdots, \eta_n, b_1, \cdots, b_n)$ is a specialization of $(\eta_1, \cdots, \eta_n, \tau\eta_1, \cdots, \tau\eta_n)$ over $\mathfrak{F}$, that is, $(b_1, \cdots, b_n)$ is a specialization of $(\tau\eta_1, \cdots, \tau\eta_n)$ *over* $\mathfrak{G}$. Therefore if we set

$$g_i = (R_{i1}(\eta_1, \cdots, \eta_n, b_1, \cdots, b_n), \cdots, R_{id_i}(\eta_1, \cdots, \eta_n, b_1, \cdots, b_n))$$

then $(g_i, \sigma g_i)$ is a specialization of $(f_i(\tau), \sigma f(\sigma^{-1}\tau))$ over $\mathfrak{G}$ for every i. Accordingly, the affine points g_i are representatives of an abstract point g which is rational over $\mathfrak{G}$, and $(g, \sigma g)$ is a specialization of $(f(\tau), \sigma f(\sigma^{-1}\tau))$ over $\mathfrak{G}$. But $\tau = \sigma \cdot \sigma^{-1}\tau$, so that $f(\tau) = f(\sigma)\sigma f(\sigma^{-1}\tau)$, and $f(\sigma)$ is rational over $\mathfrak{G}$. Therefore $g = f(\sigma)\sigma g$ for all $\sigma \in \mathfrak{S}$ which are automorphisms of $\mathfrak{G}$, hence for all $\sigma \in \mathfrak{S}$; that is, $f \in B^1(\mathfrak{S}, G)$.

Proof of 3). Let $\mathfrak{F}'$ be the algebraic closure of $\mathfrak{F}$. $\mathfrak{F}'$ is a normal algebraic extension of $\mathfrak{F}$, and $\mathfrak{F}' \cap \mathfrak{G}$ is the relative algebraic closure $\mathfrak{F}^0$ of $\mathfrak{F}$ in $\mathfrak{G}$. Therefore every automorphism τ of $\mathfrak{F}'$ over $\mathfrak{F}^0$ can be extended to a unique automorphism (which we also denote by τ) of the compositum $\mathfrak{F}'\langle\mathfrak{G}\rangle = \mathfrak{G}\langle\mathfrak{F}'\rangle$ over $\mathfrak{G}$ (the group of automorphisms of $\mathfrak{F}'$ over $\mathfrak{F}^0$ thereby becoming identified with the group of automorphisms of $\mathfrak{G}\langle\mathfrak{F}'\rangle$ over $\mathfrak{G}$); also ([7], Ch. III, § 3, Th. 5, p. 799) $\mathfrak{G}\langle\mathfrak{F}'\rangle$ is a strongly normal extension of $\mathfrak{F}'$, and every isomorphism σ of $\mathfrak{G}$ over $\mathfrak{F}^0$ can be extended to a unique isomorphism (which we also denote by σ) of $\mathfrak{G}\langle\mathfrak{F}'\rangle$ over $\mathfrak{F}'$ (the component of identity $\mathfrak{S}^0$ of $\mathfrak{S}$ thereby becoming identified with the Galois group of $\mathfrak{G}\langle\mathfrak{F}'\rangle$ over $\mathfrak{F}'$). Since every element of $\mathfrak{G}\langle\mathfrak{F}'\rangle$ can be expressed as the quotient of two elements of the form $\sum \psi_i \phi_i$ $(\psi_i \varepsilon \mathcal{G}, \phi_i \varepsilon \mathcal{F}')$, and since when σ is an *auto*morphism

$$\sigma\tau(\sum \psi_i\phi_i) = \sigma \sum \psi_i \tau\phi_i = \sum (\sigma\psi_i)\tau\phi_i = \tau \sum (\sigma\psi_i)\phi_i = \tau\sigma(\sum \psi_i\phi_i),$$

we see that $\sigma\tau = \tau\sigma$. It follows that if $\alpha_1, \cdots, \alpha_h$ are the conjugates of $\alpha \varepsilon \mathcal{G}\langle\mathcal{F}'\rangle$ over $\mathcal{G}$ then $\sigma\alpha_1, \cdots, \sigma\alpha_h$ are the conjugates of $\sigma\alpha$ over $\mathcal{G}$ (where σ is any element of $\mathfrak{G}^0$ which is an automorphism of $\mathcal{G}$).

Now let $f \varepsilon Z^1(\mathfrak{G}, G)$. Since $\mathfrak{G}^0$ can be identified with the Galois group of $\mathcal{G}\langle\mathcal{F}'\rangle$ over $\mathcal{F}'$, since (for $\sigma \varepsilon \mathfrak{G}^0$) $\mathcal{G}\langle\mathcal{F}'\rangle(f(\sigma)) \subset \mathcal{G}\langle\mathcal{F}'\rangle(\sigma(\mathcal{G}\langle\mathcal{F}'\rangle))$, and since (for $\sigma, \sigma' \varepsilon \mathfrak{G}^0$) "$\sigma'$ is a specialization of σ (as isomorphism of $\mathcal{G}\langle\mathcal{F}'\rangle$)" implies "$\sigma'$ is a specialization of σ (as isomorphisms of $\mathcal{G}$)" which implies "$f(\sigma')$ is a specialization of $f(\sigma)$ over $\mathcal{G}$" which (because $\mathcal{G}\langle\mathcal{C}^*\rangle$ and $\mathcal{G}\langle\mathcal{F}'\rangle$ are linearly disjoint over $\mathcal{G}$) in turn implies "$f(\sigma')$ is a specialization of $f(\sigma)$ over $\mathcal{G}\langle\mathcal{F}'\rangle$," we see that the restriction of f to $\mathfrak{G}^0$ may be considered as a rational crossed homomorphism of the Galois group of $\mathcal{G}\langle\mathcal{F}'\rangle$ over $\mathcal{F}'$ into G. By 2) above, therefore, there exists an $\alpha \varepsilon \mathfrak{G}^\dagger$ rational over $\mathcal{G}\langle\mathcal{F}'\rangle$, such that $f(\sigma) = \alpha^{-1}\sigma\alpha$ for every $\sigma \varepsilon \mathfrak{G}^0$ which is an automorphism of $\mathcal{G}$. Denoting the conjugates of α over $\mathcal{G}$ by $\alpha_1, \cdots, \alpha_h$ we see by the above (since $f(\sigma)$ is rational over $\mathcal{G}$) that $f(\sigma) = \alpha_i^{-1}\sigma\alpha_i$ $(1 \leqq i \leqq h)$, Since G is by assumption commutative, we theerfore have $(f(\sigma))^h = \beta^{-1}\sigma\beta$, where $\beta = \alpha_1 \cdots \alpha_h$ is rational over $\mathcal{G}$. Therefore the 1-cocycle g cohomologous to f^h, defined by $g(\sigma) = \beta(f(\sigma))^h\sigma\beta^{-1}$ $(\sigma \varepsilon \mathfrak{G})$, has the value 1 at every element of $\mathfrak{G}^0$. It follows, as in the proof of 1) above, that g is constant on each coset and that $g(\sigma)$ is rational over $\mathcal{F}^0$ for every $\sigma \varepsilon \mathfrak{G}$. Consequently, letting $\sigma_1, \cdots, \sigma_k$ denote a complete set of representatives of the cosets of $\mathfrak{G}$ relative to $\mathfrak{G}^0$, and setting $\gamma = g(\sigma_1) \cdots g(\sigma_k)$, we find that

$$\sigma\gamma = \prod \sigma g(\sigma_j) = \prod g(\sigma)^{-1} g(\sigma\sigma_j) = (g(\sigma))^{-k}\gamma$$

for every $\sigma \varepsilon \mathfrak{G}$, so that $(f(\sigma))^{hk} = \beta^{-k}\gamma\sigma(\gamma^{-1}\beta^k)$ for every $\sigma \varepsilon \mathfrak{G}$ and $f^{hk} \varepsilon B^1(\mathfrak{G}, G)$.

3. Applications. Theorem 1 of Chapter II asserts that there exists a birational isomorphism of the Galois group $\mathfrak{G}$ onto a group variety G. Now such a birational isomorphism, or more generally any rational homomorphism f of $\mathfrak{G}$ into G, is a 1-cocycle of $\mathfrak{G}$ into G. If we can prove that f is a 1-coboundary then there exists a point α of G (or, more precisely, of $G^\dagger$) which is rational over $\mathcal{G}$ such that $f(\sigma) = \alpha^{-1}\sigma\alpha$, that is, such that $\sigma\alpha = \alpha f(\sigma)$ for every $\sigma \varepsilon \mathfrak{G}$. In the case in which f is actually an isomorphism, we see that α is invariant under σ only when σ is the identity isomorphism, so that $\mathcal{G} = \mathcal{F}\langle\alpha\rangle$. More generally, if f is a homomorphism with *finite* kernel then $\mathcal{G}$ is a normal algebraic extension of $\mathcal{F}\langle\alpha\rangle$ with Galois group isomorphic to the kernel of f. These remarks are the basis of the use of Theorem 1 in the proofs of the following two theorems.

THEOREM 2. *A necessary and sufficient condition that the strongly normal extension $\mathcal{G}$ of $\mathcal{F}$ be a Picard-Vessiot extension is that its Galois group $\mathfrak{G}$ be birationally isomorphic to an algebraic group of matrices.*

Proof. If $\mathcal{G}$ is a Picard-Vessiot extension of $\mathcal{F}$ then, as we have already observed (Chapter I, § 3, Example 2), $\mathfrak{G}$ is birationally isomorphic to an algebraic matric group. Conversely, let f be a birational isomorphism of $\mathfrak{G}$ onto an algebraic subgroup of $GL(n)$. By Theorem 1, there exists an invertible square matrix $\alpha = (\alpha_{ij})_{1 \leqq i \leqq n, 1 \leqq j \leqq n}$ with coordinates $\alpha_{ij} \in \mathcal{G}$ such that $\sigma\alpha = \alpha f(\sigma)$ for every $\sigma \in \mathfrak{G}$, and $\mathcal{G} = \mathcal{F}\langle\alpha\rangle$. Therefore if $\eta_1, \cdots, \eta_p$ form a maximal subset of the set of elements α_{ij} linearly independent over $\mathcal{K}$ (and therefore over every field of constants) then for each $\sigma \in \mathfrak{G}$ there exists an invertible matrix $(c_{ij}(\sigma))_{1 \leqq i \leqq p, 1 \leqq j \leqq p}$ with constant coordinates $c_{ij}(\sigma)$, such that $\sigma\eta_j = \sum_{1 \leqq i \leqq p} c_{ij}(\sigma)\eta_i$ $(1 \leqq j \leqq p)$. It is now clear that if $\theta_1, \cdots, \theta_p$ are arbitrary operators of the form $\delta_1^{i_1} \cdots \delta_m^{i_m}$ then the determinant

$$W_{\theta_1 \cdots \theta_p}(\eta_1, \cdots, \eta_p) = \det(\theta_i\eta_j)_{1 \leqq i \leqq p, 1 \leqq j \leqq p}$$

has the property that

$$\sigma W_{\theta_1 \cdots \theta_p}(\eta_1, \cdots, \eta_p) = \det(c_{ij}(\sigma)) \cdot W_{\theta_1 \cdots \theta_p}(\eta_1, \cdots, \eta_p),$$

so that the ratio of any two nonzero determinants of this type is invariant under $\mathfrak{G}$ and hence is an element of $\mathcal{F}$. Thus ([6], § 3, p. 599) $\mathcal{G}$ is a Picard-Vessiot extension of $\mathcal{F}$.

COROLLARY. *If $\mathcal{G}$ is a Picard-Vessiot extension of $\mathcal{F}$ and $\mathcal{M}$ is any differential field between $\mathcal{F}$ and $\mathcal{G}$ which is a strongly normal extension*[3] *of $\mathcal{F}$ then $\mathcal{M}$ is a Picard-Vessiot extension of $\mathcal{F}$.*

Proof. Since $\mathcal{G}$ is Picard-Vessiot over $\mathcal{F}$ there exists a birational isomorphism f of $\mathfrak{G}$ onto an algebraic matric group G. The Galois group $\mathfrak{G}(\mathcal{M})$ of $\mathcal{G}$ over $\mathcal{M}$ is a normal algebraic subgroup of $\mathfrak{G}$, therefore $K = f(\mathfrak{G}(\mathcal{M}))$ is a normal algebraic subgroup of G. Now it is known (Chevalley and Kolchin [4], § 1, Th. 3) that there exists a matric representation g of G with kernel K; furthermore, although it is not specifically stated in [4], the proof there (§ 4) provides such a g which is a rational homomorphism of G. If we denote the image of g by H then (Chapter I, corollary to Proposition 5) H is an algebraic matric group. Clearly $g \circ f$ is a rational homomorphism of $\mathfrak{G}$ with kernel $\mathfrak{G}(\mathcal{M})$ and image H. Now (Chapter I, § 3, Example 1) the mapping r of $\mathfrak{G}$, which to each $\sigma \in \mathfrak{G}$ corresponds the restriction of σ

---

[3] Or even ([7], Ch. III, § 3, Th. 4, p. 797) a weakly normal extension.

to $\mathcal{M}$, is a rational homomorphism of $\mathfrak{G}$ with kernel $\mathfrak{G}(\mathcal{M})$ and image the Galois group $\mathfrak{H}$ of $\mathcal{M}$ over $\mathcal{F}$. Therefore (Chapter II, § 3, Proposition 1) there exists a birational isomorphism of $\mathfrak{H}$ onto H so that, by the present theorem, $\mathcal{M}$ is a Picard-Vessiot extension of $\mathcal{F}$.

THEOREM 3. *If the strongly normal extension $\mathcal{G}$ of $\mathcal{F}$ is of transcendence degree* 1 *and if $\mathcal{F}$ is relatively algebraically closed in $\mathcal{G}$ then there exists an element $\alpha \in \mathcal{G}$ such that either α is primitive over $\mathcal{F}$ and $\mathcal{G} = \mathcal{F}\langle\alpha\rangle$, or α is exponential over $\mathcal{F}$ and $\mathcal{G} = \mathcal{F}\langle\alpha\rangle$, or α is weierstrassian over $\mathcal{F}$ and $\mathcal{G}$ is an abelian algebraic extension of $\mathcal{F}\langle\alpha\rangle$ of finite degree. In the last case, if $\mathcal{F}$ is algebraically closed then the weierstrassian element α may be chosen so that $\mathcal{G} = \mathcal{F}\langle\alpha\rangle$.*

This theorem is not new, being the main substance of Th. 7 of [7], Ch. III, p. 809. We include it here to show how it can be derived as a rather easy consequence of Theorem 1, and Theorem 1 of Chapter II; this makes it possible to avoid the long, tedious, and intricate computations of the proof in [7], pp. 809-823.

Proof. It follows from the hypothesis that the Galois group $\mathfrak{G}$ has dimension 1 and is irreducible; by Theorem 1 of Chapter II there exists a birational isomorphism c of $\mathfrak{G}$ onto a group variety G defined over $\mathcal{K}$, which must also be of dimension 1 and irreducible. It is known (see Weil [14], § V, No. 38, pp. 68-69) that every such group variety is birationally isomorphic to (and therefore we may assume equal to) either the additive group of the field $\mathcal{K}^*$ of all constants or the multiplicative group of $\mathcal{K}^*$, or else a group variety which is a complete curve of genus 1 in projective space. In the last case we may always assume (for example, see Chevalley [2], Ch. II, §3) that the curve has an equation of the form $X_0X_2{}^2 = P(X_0, X_1)$, where P is a homogenous polynomial in $\mathcal{K}[X_0, X_1]$ of degree 3 without multiple factor; applying a suitable transformation $(X_0, X_1, X_2) \to (X_0, aX_0 + bX_1, cX_2)$ if necessary, we lose no generality in supposing that

$$P(X_0, X_1) = 4X_1{}^3 - g_2X_0{}^2X_1 - g_3X_0{}^3,$$

where $g_2, g_3 \in \mathcal{K}$. We may take the law of composition so that the group unity is $(0:0:1)$, which is the sole point at infinity on the curve, and such that the product $(1:\eta:\zeta)$ of two points $(1:\alpha:\beta)$ and $(1:a:b)$ with $\alpha \neq a$ is given by

$$\eta = -\alpha - a + 4^{-1}(\alpha - a)^{-2}(\beta - b)^2, \tag{3}$$

$$\zeta = -2^{-1}(\beta + b) + 2^{-1}3(\alpha - a)^{-1}(\beta - b) - 4^{-1}(\alpha - a)^{-3}(\beta - b)^3. \tag{4}$$

Consider the first case. By Theorem 1, there exists an element $\alpha \in \mathcal{G}$ such that $c(\sigma) = \sigma\alpha - \alpha$ for every $\sigma \in \mathfrak{G}$. Since $c(\sigma) = 0$ only when σ is the identity, only the identity in $\mathfrak{G}$ leaves α invariant, so that $\mathcal{G} = \mathcal{F}\langle\alpha\rangle$. Applying the derivation operator δ_i to both sides of the equation $\sigma\alpha = \alpha + c(\sigma)$ we obtain $\sigma\delta_i\alpha = \delta_i\alpha (\sigma \in \mathfrak{G})$, so that $\delta_i\alpha \in \mathcal{F}$ $(1 \leqq i \leqq m)$, and α is primitive over $\mathcal{F}$.

Next consider the second case. By Theorem 1, there exists a nonzero element $\alpha \in \mathcal{G}$ such that $c(\sigma) = \alpha^{-1}\sigma\alpha$ for every $\sigma \in \mathfrak{G}$. Since $c(\sigma) = 1$ only when σ is the identity, only the identity in $\mathfrak{G}$ leaves α invariant, so that $\mathcal{G} = \mathcal{F}\langle\alpha\rangle$. From the equation $\sigma\alpha = \alpha c(\sigma)$ we find that $\sigma\delta_i\alpha = (\delta_i\alpha)c(\sigma)$, whence $\sigma(\alpha^{-1}\delta_i\alpha) = \alpha^{-1}\delta_i\alpha$ $(\sigma \in \mathfrak{G})$, so that $\alpha^{-1}\delta_i\alpha \in \mathcal{F}$ $(1 \leqq i \leqq m)$, and α is exponential over $\mathcal{F}$.

Consider finally the third case. By Theorem 1, there exists a point $(1:\alpha:\beta)$ on the curve $G^{\dagger}$, with $\alpha, \beta \in \mathcal{G}$, and there exists an integer $n > 0$ ($n = 1$ if $\mathcal{F}$ is algebraically closed), such that $c(\sigma)^n = (1:\alpha:\beta)^{-1}\,(1:\sigma\alpha:\sigma\beta)$ for every $\sigma \in \mathfrak{G}$. Not every element of $\mathfrak{G}$ has order n, so that $c(\sigma)^n$ fails to be the identity for some $\sigma \in \mathfrak{G}$; therefore $(1:\alpha:\beta)$ is not an invariant of the whole group $\mathfrak{G}$, hence is not rational over $\mathcal{F}$. Since $\mathcal{F}$ is relatively algebraically closed in $\mathcal{G}$ and $\mathcal{G}$ is of transcendence degree 1 over $\mathcal{F}$, $\mathcal{G}$ is an algebraic extension of $\mathcal{F}\langle\alpha, \beta\rangle$ which clearly must be of finite degree, normal, and indeed abelian; if $\mathcal{F}$ is algebraically closed, so that $n = 1$, then $\mathcal{G} = \mathcal{F}\langle\alpha, \beta\rangle$. For any $\sigma \in \mathfrak{G}$ which is not of order n (in particular, for a generic element of $\mathfrak{G}$) we may write $c(\sigma)^n = (1:a:b)$, where a, b are constants. Then $\sigma\alpha$ and $\sigma\beta$ are given by the second members of (3) and (4), respectively. Making use of the fact that

$$\beta^2 = 4\alpha^3 - g_2\alpha - g_3, \tag{5}$$

these may be rewritten to obtain

$$\sigma\alpha = 4^{-1}(\alpha - a)^{-2}[4a\alpha^2 + (4a^2 - g_2)\alpha - (g_2 a + 2g_3) - 2b\beta], \tag{6}$$

$$\begin{aligned} \sigma\beta = 4^{-1}(\alpha - a)^{-3}[&(4\alpha^3 + 12a\alpha^2 - 3g_2\alpha - g_2 a - 4g_3)b \\ &+ ((-12a^2 + g_2)\alpha - 4a^3 + 3g_2 a + 4g_3)\beta]. \end{aligned} \tag{7}$$

From (5) we find $2\beta\delta_i\beta = (12\alpha^2 - g_2)\delta_i\alpha$. Therefore application of δ_i to (6) yields

$$\begin{aligned} \sigma\delta_i\alpha &= 4^{-1}(\alpha - a)^{-2}((8a\alpha + 4a^2 - g_2)\delta_i\alpha - 2b)\delta_i\beta \\ &\quad - 2^{-1}(\alpha - a)^{-3}(4a\alpha^2 + (4a^2 - g_2)\alpha - g_2 a - 2g_3 - 2b\beta)\delta_i\alpha \\ &= 4^{-1}(\alpha - a)^{-2}((8a\alpha + 4a^2 - g_2)\beta - b(12\alpha^2 - g_2))\beta^{-1}\delta_i\alpha \\ &\quad - 2^{-1}(\alpha - a)^{-3}((4a\alpha^2 + (4a^2 - g_2)\alpha - g_2 a - 2g_3)\beta \cdot \\ &\qquad - 2b(4\alpha^3 - g_2\alpha - g_3))\beta^{-1}\delta_i\alpha. \end{aligned}$$

The last member equals $\beta^{-1}\delta_i\alpha$ multiplied by the second member of (7). Therefore $\sigma\delta_i\alpha = (\sigma\beta)\beta^{-1}\delta_i\alpha$, that is $\sigma(\beta^{-1}\delta_i\alpha) = \beta^{-1}\delta_i\alpha$ $(\sigma \varepsilon \mathfrak{G})$, so that $\beta^{-1}\delta_i\alpha \varepsilon \mathcal{F}$ $(1 \leqq i \leqq m)$. It follows that $\mathcal{F}\langle\alpha,\beta\rangle = \mathcal{F}\langle\alpha\rangle$, and also (with the help of (5) that $(4\alpha^3 - g_2\alpha - g_3)^{-1}$ $(\delta_i\alpha)^2$ is the square of an element of $\mathcal{F}$ $(1 \leqq i \leqq m)$, that is, that α is weierstrassian over $\mathcal{F}$.

COLUMBIA UNIVERSITY
AND UNIVERSITY OF PARIS.

REFERENCES.

[1] N. Bourbaki, *"Algèbre,"* Chapters IV-V (Actualités scientifiques et industrielles 1102), Hermann et Cie., Paris, 1950.

[2] C. Chevalley, *"Introduction to the theory of algebraic functions of one variable"* (Mathematical Surveys VI), American Mathematical Society, New York, 1951.

[3] ———, *"Théorie des groupes de Lie, Tome II: Groupes algébriques"* (Actualités scientifiques et industrielles 1152), Hermann et Cie., Paris, 1951.

[4] ———, and E. Kolchin, "Two proofs of a theorem on algebraic groups," *Proceedings of the American Mathematical Society,* vol. 2 (1951), pp. 126-134.

[5] E. R. Kolchin, "Algebraic matric groups and the Picard-Vessiot theory of homogeneous linear ordinary differential equations," *Annals of Mathematics,* vol. 49 (1948), pp. 1-42.

[6] ———, "Picard-Vessiot theory of partial differential fields," *Proceedings of the American Mathematical Society,* vol. 3 (1952), pp. 596-603.

[7] ———, "Galois theory of differential fields," *American Journal of Mathematics,* vol. 75 (1953), pp. 753-824.

[8] H. Matsumura, "Automorphism-groups of differential fields and group varieties," *Memoirs of the College of Science, University of Kyoto, Series A,* vol. 28 (1954), pp. 283-292.

[9] S. Nakano, "On invariant differential forms on a group variety," *Journal of the Mathematical Society of Japan,* vol. 2 (1951), pp. 216-227.

[10] ———, "Note on group varieties," *Memoirs of the College of Science, University of Kyoto, Series A,* vol. 27 (1952), pp. 55-66.

[11] M. Rosenlicht, "Generalized Jacobian varieties," *Annals of Mathematics,* vol. 59 (1954), pp. 505-530.

[12] J.-P. Serre, "Faisceaux algébriques cohérents," *Annals of Mathematics,* vol. 61 (1955), pp. 197-278.

[13] A. Weil, *"Foundations of Algebraic geometry"* (American Mathematical Society Colloquium Publications, vol. 29), New York, 1946.

[14] ———, "Variétés abéliennes et courbes algébrique (Actualités scientifiques et industrielles 1064), Hermann et Cie., Paris, 1948.

ENGEL RINGS AND A RESULT OF HERSTEIN AND KAPLANSKY.*

By M. P. Drazin.

1. Introduction. It is known that if, in an associative ring R, some power of every element is central, then the commutator ideal R_1 of R (i. e. the two-sided ideal of R generated by all the commutators $[a,b]=ab-ba$ with a, b in R) is nil; more precisely, if to each element y of R corresponds a positive integer $n=n(y)$ such that $[x,y^n]=0$ for all x in R, then to each element c of R_1 corresponds a positive integer j such that $c^j=0$. Special cases of this were established by Jacobson, Hua and others, after which Kaplansky [15] showed in the general case that R_1 at any rate lies in the Jacobson radical; that R_1 is actually nil was finally proved by Herstein [9]. It will be convenient to call a given ring R *commutator-nil* whenever its commutator ideal R_1 is nil.

Consider now another class of rings. For any x, y in R, define a sequence of elements $e_k(x,y)$ of R inductively by the rules

$$e_0(x,y)=x, \qquad e_k(x,y)=[e_{k-1}(x,y),y] \qquad (k=1,2,\cdots);$$

equivalently, in binomial notation,

$$e_k(x,y)=\sum_{j=0}^{k}(-1)^j C_j^k y^j x y^{k-j} \qquad (k=0,1,\cdots).$$

If to each pair x, y there corresponds an integer $k=k(x,y)$ such that $e_k(x,y)=0$, then we call R an *Engel ring*. The present paper grew from an investigation of the properties of the commutator ideals of arbitrary Engel rings. The relevance of the above-mentioned result of Herstein and Kaplansky soon appeared; indeed, it became clear that the two kinds of rings can profitably be regarded as special cases of another type, the *K-rings*, where we say that a given ring R is a K-ring if, to each pair of elements x, y of R, there correspond positive integers $k=k(x,y)$ and $n=n(x,y)$ such that $e_k(x,y^n)=0$. The rings studied by Herstein and Kaplansky are obtained on imposing the restrictions that $k=1$ for all x,y and that n be independent of x, while of course the Engel rings are just those K-rings with $n=1$ for all x,y.

---

* Received June 9, 1955.

Two of our principal results (both generalizing Herstein's) are that a given K-ring must always be commutator-nil whenever its index functions k, n are both functions of y alone, and whenever $k=1$ for all x, y. Since every nil ring is a K-ring (and indeed falls in all of the special classes of K-rings we have mentioned), to prove without further hypothesis that R_1 is more than merely nil would imply a non-trivial statement about the commutator-ideal of an arbitrary nil ring. We establish no such statement here, being instead more concerned in our later results with extending the commutator-nil property (or even slightly weaker properties) to wider classes of rings. For example, we prove that the commutator ideal of an arbitrary K-ring is always contained in the Jacobson radical.

For the proofs of the facts, we need two auxiliary results about the structure of (the semi-simple part of) an arbitrary K-ring. These are derived in a more general context in Section 2. We define, for each positive integer m, a class of rings called m-rings. These classes form an expanding sequence, i. e., if l, m are any given integers satisfying $1 \leqq l \leqq m$, then every l-ring is an m-ring. Every K-ring is a 1-ring (and hence an m-ring for all m); also every algebraic ring of bounded index, and every ring with minimal condition on its left (or right) ideals is an m-ring for some m.

Four successively narrower classes of rings have now been mentioned: the m-rings (m arbitrary), the 1-rings, the K-rings and the Engel rings. Having discussed the first and widest class in Section 2 we proceed step by step from the general to the particular, examining the other three classes in turn in Sections 3, 4 and 5. In Section 3 we consider the 1-rings (which include, besides the K-rings, all strongly regular rings and another class of rings investigated recently by Herstein [10] and others), restating the specialization $m=1$ of the results of Section 2, and proving also one other result. Our theorems on K-rings as such are proved in Section 4. In Section 5 we press our investigation a little further for the Engel rings, showing that certain of these satisfy conditions stronger than the commutator-nil property.

In our final section we consider those rings in which, for each pair of elements x, y, there are integers k, m, n such that $e_k(x^m, y^n)=0$. These include all the K-rings as the special case $m=1$, and we show that many of our results for K-rings can be extended to this more general class.

Throughout, R will denote an associative ring, not necessarily with a unit element, admitting a commutative ring F of operators (i. e. endomorphisms α of the additive group of R, subject to $\alpha(xy)=(\alpha x)y=x(\alpha y)$ for all $x, y \varepsilon R$); we may suppose without loss of generality that F contains the identity operator. The case of a "ring without operators" is included in this scheme on taking F to be just the ring of integers (or an appropriate quotient ring). When we refer to subrings (etc.) of R these should always be understood as sub-F-rings (etc.), i.e. as being mapped into themselves by every operator in F.

2. Two theorems about *m*-rings. Given any ring R over F and any positive integer m, we shall say that R is an *m-ring* whenever the following statement is true: to each choice of x, y in R there correspond $\alpha=\alpha(x,y)$ in F, $a=a(x,y)$ in R and integers $k=k(x,y)$, $l=l(x,y)$, $n=n(x,y)$ such that

$$1\leqq l\leqq m, \qquad x^{m-l}e_k(\alpha x^{l+1}+x^{l+1}a-x^l,y^n)=0.$$

Every algebraic ring of bounded index, and every ring with minimal condition on its left or right ideals, is of this type for some m (even with $k=0$, $l=m$, $n=1$). We establish now two results about the structure of the semi-simple parts of arbitrary m-rings, interpreting semi-simplicity in the sense of Koethe for Theorem 2.1, and in the sense of Jacobson for Theorem 2.2.

THEOREM 2.1. *Let R be a given m-ring without non-zero nil two-sided ideals. Then every nilpotent element z of R satisfies $z^{m^2}=0$.*

Proof. The main step of the proof is to show that if a given element x of R satisfies $x^{m+1}=0$, then x^m generates a nil two-sided ideal (and must consequently vanish).

We show first that, for any $b\in R$, bx^m is nilpotent. We have, for any $\alpha\in F$, any $a\in R$ and any positive integers k, l, n with $l\leqq m$,

$$\begin{aligned}&x^{m-l}e_k(\alpha x^{l+1}+x^{l+1}a-x^l,(bx^m)^n)\\&\qquad=x^{m-l}(\alpha x^{l+1}+x^{l+1}a-x^l)(bx^m)^{nk}=-x^m(bx^m)^{nk};\end{aligned}$$

but R is an m-ring, so this, and consequently $(bx^m)^{nk+1}$, vanishes for some choice of k,n.

We can now deduce that x^m generates a nil two-sided ideal. The typical element of this ideal has the form $d=\beta x^m+bx^m+x^mc$, where $\beta\in F$ and $b,c\in R$. Using the fact that $bx^m\cdot x^m=x^m\cdot x^mc=0$, we see that, for each positive integer j,

$$\begin{aligned}d^j&=\beta\sum_{i=0}^{j-1}(x^mc)^ix^m(bx^m)^{j-i-1}+\sum_{i=0}^{j}(x^mc)^i(bx^m)^{j-i}\\&=\beta x^m\sum_{i=0}^{j-1}(cx^m)^i(bx^m)^{j-i-1}+(bx^m)^j+x^m\sum_{i=1}^{j}(cx^m)^{i-1}c(bx^m)^{j-i}.\end{aligned}$$

Also, by what we have already proved, there are positive integers p, q such that $(bx^m)^p = (cx^m)^q = 0$, whence it follows that $d^{p+q} = 0$.

We have now shown that $x^{m+1} = 0$ implies $x^m = 0$. Let z be an arbitrary nilpotent element of R, and let r be the smallest positive integer for which $z^{rm} = 0$. Then, if $r > m$, we should have $(z^{r-1})^{m+1} = z^{r-m-1}z^{rm} = 0$, whence, on taking $x = z^{r-1}$, it would follow that $(z^{r-1})^m = 0$, contrary to our choice of r. Consequently $r \leqq m$, i.e. $z^{m^2} = 0$ as required.

We cannot, even for the 1-rings, hope to replace the hypothesis of no nil ideals in Theorem 2.1 by the weaker one of no nilpotent ideals; for Baer [3; 540-543] has constructed a non-zero nil ring having no non-zero nilpotent ideals. Indeed, by Nagata [18] there exist nil rings (and *a fortiori* 1-rings and m-rings), having no non-zero nilpotent ideals, whose nilpotent elements are of unbounded index.

We have stated Theorem 2.1 in a way which avoids explicit reference to the Koethe radical (i.e. the union of all two-sided nil ideals of R). An equivalent statement is the following: If R is any given m-ring, with Koethe radical N, say, then, given any $z \varepsilon R$, the existence of an integer s with $z^s \varepsilon N$ implies $z^{m^2} \varepsilon N$ (in particular, R/N has no non-zero nilpotent elements if $m = 1$). If J denotes the Jacobson radical of R, then an analogous statement holds with N replaced by J (since R/J has no non-zero nil ideals).

There being regrettably few problems which yield to attacks based on the rather weak structure theory associated with the Koethe radical, we next prove another auxiliary theorem for use when the Jacobson theory is appropriate. This generalizes a result of Arens and Kaplansky [2; Theorem 3.1]; to describe their result in our terminology, they considered only m-rings in which $k = 0$ for all x, y.

THEOREM 2.2. *Let R be any given primitive m-ring over F. Then R is isomorphic to the complete ring $M_q(D)$ of $q \times q$ matrices over D, where $q \leqq m$ and D is a suitable division ring over F.*

Proof. R, being primitive, is isomorphic to a dense ring M of linear transformations of a vector space V over some division ring D; we shall denote the result of operating on $v \varepsilon V$ with $x \varepsilon M$ by vx, i.e. regard M as operating on V from the right.[1] We have only to show that V cannot contain $m + 1$ elements $v_0, v_1, \cdots, v_m$ independent with respect to D.

In the contrary case, since M is dense, we could choose x, y in M so that

---

[1] This is dictated by our having used the expression $\alpha x^{l+1} + x^{l+1}a - x^l$ rather than $\alpha x^{l+1} + ax^{l+1} - x^l$ when defining m-rings (so as to ensure that every strongly regular ring is self-evidently a 1-ring).

$$v_h x = v_{h+1} \qquad (h = 0, 1, \cdots, m-1), \qquad v_m x = 0,$$

$$v_h y = 0 \qquad (h = 0, 1, \cdots, m-1), \qquad v_m y = v_m.$$

Then, for any $\alpha \varepsilon F$, any $a \varepsilon R$ and any $k \geqq 0$, $1 \leqq l \leqq m$, $n \geqq 1$,

$$v_0 x^{m-l} e_k(\alpha x^{l+1} + x^{l+1}a - x^l, y^n) = v_{m-l} e_k(\alpha x^{l+1} + x^{l+1}a - x^l, y^n)$$

$$= v_{m-l}(\alpha x^{l+1} + x^{l+1}a - x^l)(y^n)^k$$

$$= v_{m-l} x^{l+1}(\alpha y^{nk} + a y^{nk}) - v_{m-l} x^l y^{nk} = 0 - v_m y^{nk} = -v_m \neq 0$$

(by the D-independence of $v_0, v_1, \cdots, v_m$), contrary to the m-ring hypothesis.

Remark. The briefest examination of the proof of Theorem 2.1 shows that the full force of the m-ring hypothesis is not needed, but only its consequence that, for any b, x in R, $x^{m+1} = 0$ implies bx^m nilpotent. This will always be the case whenever, for each choice of b, x in R, a positive integer s exists such that $(bx^m)^s$ is in the two-sided ideal generated by x^{m+1}; and this is in turn entailed (on taking $y = bx^m$) by the hypothesis that, for each choice of x, y in R, an integer $t = t(x, y)$ exists such that $x^m y^t$ is in the two-sided ideal generated by yx and x^{m+1} (which condition is still satisfied by every m-ring, with $t(x, y) = k(x, y) n(x, y)$). These trivial comments are relevant to certain generalizations of Theorem 2.2 which will be described in detail elsewhere [6].

3. Remarks on the case $m = 1$.

To say that R is a 1-ring means that, to each choice of x, y in R, there correspond α in F, a in R and integers k, n such that $e_k(\alpha x^2 + x^2 a - x, y^n) = 0$. It is not hard to see that, by replacing k and n respectively by

$$\max\{k(x, y), k(\alpha(x, y)x^2 + x^2 a(x, y) - x, y)\}$$

and

$$n(x, y) n(\alpha(x, y)x^2 + x^2 a(x, y) - x, y),$$

and by modifying a appropriately, we may always take $\alpha = 0$ in a 1-ring.

Applying Theorem 2.1 to R/N with $m = 1$, we have at once

THEOREM 3.1. *In any 1-ring, every nilpotent element generates a nil two-sided ideal.*

Analogously, it is easy to see that, in any commutator-nil ring, every nilpotent element generates a nil two-sided ideal. However, not every 1-ring is commutator-nil (consider for example any non-commutative division ring); and one can also verify without difficulty (e.g. from Theorem 3.3 below) that not every commutator-nil ring is a 1-ring.

Next, for any ring R, with Jacobson radical J, R/J is semi-simple in Jacobson's sense, so that R/J is isomorphic to a subdirect sum of primitive rings, each of which is a homomorphic image of R/J and hence of R. Consequently, by Theorem 2.2 and the homomorphism-invariance of the 1-ring property, we have

THEOREM 3.2. *If R is any 1-ring, with Jacobson radical J, then R/J is a subdirect sum of division rings.*

This may be regarded as a generalization of a familiar property of strongly regular rings (which have $J = 0$); it also generalizes an argument of Herstein [10; Theorem 5].

Our final result in this section is included chiefly for the sake of its connexion with the existence of annihilators of ideals in rings with minimal condition on left (or right) ideals (cf. [5]). Of course every K-ring is a 1-ring (with $\alpha = 0$, $a = 0$), and clearly every idempotent element of a K-ring is central; for if $y^2 = y$ then, for any x, n,

$$e_3(x, y^n) = e_3(x, y) = xy^3 - 3yxy^2 + 3y^2xy - y^3x = xy - yx,$$

so that $xy - yx = e_{2j+1}(x, y^n)$ $(j, n = 1, 2, \cdots)$. We show (cf. Herstein [10; Lemma 4]) that this result in fact holds for all 1-rings:

THEOREM 3.3. *In any 1-ring, every idempotent element is central.*

Proof. Let R be a given 1-ring and y any idempotent in R; we must show that, for any given $b \varepsilon R$, $by = yb$.

Define $x = by - yby$, so that $yx = 0$ and $xy = x$, whence $x^2 = 0$ and consequently (since R is a 1-ring) integers k, n exist satisfying

$$0 = e_k(-x, y^n) = -e_k(x, y) = -\sum_{j=0}^{k} (-1)^j C_j^k y^j x y^{k-j} = -xy^k = -x,$$

i.e. $yby = by$; and a similar argument shows that $yby = yb$, as required.

The above proof requires only that $y^2 = y$, together either with $yx = 0$ and $xy = x$, or with $xy = 0$ and $yx = x$, implies $x = 0$. This holds in any ring R admitting a positive integral-valued function $t(x, y)$ such that, for all $x, y \varepsilon R$, (a) xy^t is in the two-sided ideal generated by yx and x^2, and (b) $y^t x$ is in the two-sided ideal generated by xy and x^2; here (a) is just the specialization $m = 1$ of the condition mentioned at the end of Section 2, while (b) (obtainable from (a) by writing all products in reverse order) is also satisfied by every 1-ring.

It is easy to deduce from Theorems 3.1 and 3.3 (by observing that every element of R/N is a linear combination of "principal idempotent elements") that every 1-ring which is also an algebraic algebra over an algebraically closed field must be commutator-nil; however, for the special case of K-rings, this result is included in Theorem 4.5 below.

4. The K-ring theorems.

The proof of our first main result is elementary, and self-contained apart from appeals to Theorem 3.1 and to the result of Herstein mentioned in the Introduction.

LEMMA 4.1. *Let R be a ring in which, to each element y, there correspond positive integers $k = k(y)$, $n = n(y)$ such that $e_k(x, y^n) = 0$ for all x in R; suppose also that R has characteristic either* (i) *zero, or* (ii) *a prime. Then R is commutator-nil.*

Proof. (i) We need only consider rings without non-zero nilpotent elements. For suppose the result were known for these, and let N denote the union of all two-sided nil ideals of R. Then (as is well known and easy to see) N is itself a two-sided nil ideal of R, and R/N has no non-zero two-sided nil ideals. Also R/N inherits the Engel hypothesis on R and, by Theorem 3.1, is consequently even a ring without non-zero nilpotent elements; we observe further that (since N is nil) R/N has zero characteristic. Hence, if the result were known for rings of zero characteristic without nilpotent elements, then R/N would be commutative, i.e. R_1 would be nil.

Thus we may assume without loss of generality that R itself has no non-zero nilpotent elements. Given $y, n(y)$, we may suppose also that $h = k(y)$ has been chosen as small as possible subject to $h \geqq 1$; we shall show that $h = 1$. In analogy with Leibniz's formula in differential calculus we have, for any non-negative integer j and any $a, b, c \in R$,

$$e_j(ab, c) = \sum_{i=0}^{j} C_i{}^j e_i(a, c) e_{j-i}(b, c);$$

in particular, for any $x \in R$,

$$e_{2h-2}(x^2, y^n) = \sum_{i=0}^{2h-2} C_i{}^{2h-2} e_i(x, y^n) e_{2h-2-i}(x, y^n) = C_{h-1}{}^{2h-2} e_{h-1}{}^2(x, y^n).$$

Now, if $h > 1$, then $2h - 2 \geqq h$ and so the expression on the left here vanishes, whence (since R is of zero characteristic and contains no non-zero nilpotent elements) it would follow that $e_{h-1}(x, y^n) = 0$ for all x in R, contrary to our choice of h. Hence in fact $h = 1$, i.e. we have $[x, y^{n(y)}] = 0$ for all x, y in R; the result now follows from Herstein [9].

(ii) Suppose that, for some prime p, $pR = 0$. Since the binomial coefficients $C_j{}^{p^r}$ $(j = 1, \cdots, p^r - 1; r = 1, 2, \cdots)$ are all divisible by p, we have, for all $x, z \in R$ and every positive integer r,

$$e_{p^r}(x,z) = \sum_{j=0}^{p^r} (-1)^j C_j^{p^r} z^j x z^{p^r-j} = xz^{p^r} + (-1)^{p^r} z^{p^r} x = [x, z^{p^r}]$$

(on considering separately the cases $p=2$, $p \neq 2$). Also, by our hypothesis on R, $e_{p^r}(x,y^n) = 0$ whenever $p^r \geqq k(y)$; thus, taking $z = y^n$, we see that, for all sufficiently large r (the criterion of sufficiency being independent of x), $[x, y^{n(y)p^r}] = 0$ for all $x \varepsilon R$. The result now follows from Herstein's as before.

We have next to remove the restrictions (i) and (ii) from Lemma 4.1. This is of course immediate when F is a field, and will now be done in the general case by an argument communicated to the writer by A. S. Amitsur; he naturally proved his result in the special form immediately applicable to our present purpose, but his argument seems interesting enough to deserve formal statement in somewhat more general terms. Calling an arbitrary subset Q of a ring nil whenever every element of Q is nilpotent, we have

LEMMA 4.2. *Let* $\mathfrak{R}$ *be any given set of rings, in each member* R *of which is given also a subset* $Q(R)$; *suppose further that, given any* $R \varepsilon \mathfrak{R}$ *and any homomorphism* θ *of* R, $\mathfrak{R}$ *is such that* $R\theta \varepsilon \mathfrak{R}$ *and* $Q(R)\theta \leqq Q(R\theta)$. *Then, if also* $Q(R)$ *is nil for all those* $R \varepsilon \mathfrak{R}$ *having characteristic either* (i) *zero or* (ii) *a prime,* $Q(R)$ *must in fact be nil for every* $R \varepsilon \mathfrak{R}$.

Proof. Let R be an arbitrary member of $\mathfrak{R}$, and c any element of $Q(R)$; we have to prove c nilpotent.

Let A denote the subset of R consisting of all elements of finite non-zero additive order (i.e. the "torsion subgroup" of the additive group of R). Clearly A is a two-sided ideal of R, and so R/A exists; also R/A has zero characteristic and is of course a homomorphic image of R. Hence, by (i), $Q(R/A)$ is nil, and so, by our hypothesis $Q(R)\theta \leqq Q(R\theta)$, some power c^i of c lies in A; in other words we can find positive integers i, r so that $rc^i = 0$. Let $s = s(c)$ denote the smallest positive integer annihilating a power of c, say $sc^j = 0$; we have then only to show that $s = 1$.

In the contrary case, we could find a prime p dividing s, say $s = pt$ where $1 \leqq t < s$. Since pR is a two-sided ideal of R, (ii) tells us that $Q(R/(pR))$ is nil; thus, on using once more our hypothesis $Q(R)\theta \leqq Q(R\theta)$, we see that some power of c lies in pR, say $c^k = pd$, where $d \varepsilon R$. But then

$$tc^{j+k} = ptc^j d = sc^j d = 0,$$

where $t < s$ is a positive integer, contrary to our choice of s; the result follows from this.

Now, by Lemma 4.1, on taking $Q(R) = R_1$ all the conditions of Lemma 4.2 are satisfied by the set $\mathfrak{R}$ of all rings over each of which are defined functions $k = k(y)$, $n = n(y)$ such that $e_k(x, y^n) = 0$ for all x, y. Thus we have proved

THEOREM 4.1. *If R is any ring in which, to each element y, there correspond positive integers $k=k(y)$, $n=n(y)$ such that $e_k(x,y^n)=0$ for all x, then R is commutator-nil.*

It will appear in Theorem 6.2 below that a similar result holds when it is given that $e_{k(x)}(x,y^{n(x)})=0$ for all x,y.

The proving of Theorem 4.1 has been something of a cooperative piece of work. As already stated, the argument of Lemma 4.2 is due to Amitsur, while the proof of Lemma 4.1 owes much to some (unpublished) work of P. M. Cohn and K. W. Gruenberg; the writer is glad to acknowledge his debt to these three. Cohn and Gruenberg considered only Engel rings R with $k(x,y)$ independent of x; Gruenberg had proved Theorem 3.2 for these rings, and this enabled them to prove $R_1 \leqq J$ by reduction to the division ring case. Gruenberg himself settled the subcase of prime characteristic by an argument, suggested by [12; p. 102], of which our proof of Lemma 4.1 (ii) is only a trivial extension; and Cohn, without having Theorem 3.1 to argue from, dealt with the subcase of zero characteristic by an argument rather similar to that used in the second part of our proof of Lemma 4.1 (i). For division rings these two subcases exhaust the possibilities, so that no bridging argument along the lines of Lemma 4.2 was called for. Our Theorem 4.1 improves on Cohn and Gruenberg's result in two directions, and its proof appeals to Jacobson's structure theory only indirectly (through Herstein); we extend their result in a different way in Theorem 4.4 below.

Two published results, both included in Theorem 4.1, should also be mentioned. The case of Engel rings with k independent of both x and y was settled by Amitsur [1]; indeed; he showed that in such a ring R the left (or right) ideal generated by any given element of R_1 is nil of bounded index (this was stated in his paper only for algebras over infinite fields, but he has pointed out in a letter that this restriction is inessential). His work in turn superseded a result of Jennings [14; Theorem 1], who showed that R_1 is nil whenever a fixed positive integer h exists such that the repeated commutator $[[\cdot\cdot\cdot[[x_1,x_2],x_3],\cdot\cdot\cdot],x_h]=0$ for all choices of $x_1,\cdot\cdot\cdot,x_h$ in R.

Theorem 4.1 does not fully answer the question whether every K-ring is commutator-nil; and the writer knows no counter-example to this proposition. Before discussing how the proof of Theorem 4.1 might be extended to deal with arbitrary K-rings (i.e. to allow k,n to depend also on x), we first note a slight generalization of Herstein's result:[2]

---

[2] *Added in proof.* Shortly after the present paper was submitted Herstein himself published a proof of the same generalization in the *Canadian Journal of Mathematics*, vol. 7 (1955), pp. 411-412.

THEOREM 4.2. *If in a given ring R there corresponds to each pair of elements x, y a positive integer $n=n(x,y)$ such that $xy^n=y^nx$, then R is commutator-nil.*

Proof. We observe that, since $\alpha[a,b]=[\alpha a,b]$, $\alpha d[a,b]=d[\alpha a,b]$ for any admissible operator α and any a, b, d in R, the ideal R_1 does not depend on the given set of operators on R; thus we may regard R as a ring without operators. We reduce the problem to Herstein's by showing that R is locally of the type he considered.

Let c be any element of R_1, i.e. any sum of terms of the forms $[a_i,b_i]$, $d_i[a_i,b_i]$, $[a_i,b_i]f_i$, $d_i[a_i,b_i]f_i$. Let the a_i, b_i, d_i, f_i involved in such an expression for c be enumerated as $x_1,\cdots,x_t$ (order being irrelevant), and let X denote the ring they generate. Then c is in the commutator ideal of X, and so it will be enough to prove that X is commutator-nil.

Take any $y\varepsilon R$. By the hypothesis on R, $[x_i,y^{n(x_i,y)}]=0$ $(i=1,\cdots,t)$, and so, if we define $n(y)=n(x_1,y)\cdots n(x_t,y)$, then $[x_i,y^{n(y)}]=0$ $(i=1,\cdots,t)$. Since the x_i generate X, it follows that $[x,y^{n(y)}]=0$ for all $x\varepsilon X$, $y\varepsilon R$. In particular, for each $y\varepsilon X$, $y^{n(y)}$ lies in the centre of X; hence, by Herstein's result [9], X is commutator-nil as required.

Unfortunately it does not seem possible, even when we are given that k is independent of both x and y, to remove the restriction $n=n(y)$ in Theorem 4.1 by the above sort of argument. However, using Theorem 4.2 and the argument of Lemma 4.1 (ii), we can at once deduce

THEOREM 4.3. *Every K-ring of prime characteristic is commutator-nil.*

To prove that *all* K-rings are commutator-nil (i.e. to remove the restrictions $k=k(y)$, $n=n(y)$ from Theorem 4.1) it would, in virtue of Lemma 4.2 and Theorem 4.3, be enough to consider only rings of zero characteristic, i.e. to remove the restrictions $k=k(y)$, $n=n(y)$ from Lemma 4.1 (i). And we can of course still use Theorem 3.1 to reduce further to rings without non-zero nilpotent elements (for which one would have to prove commutativity). In other words, the K-rings of zero characteristic without non-zero nilpotent elements are critical for the whole question whether all (and, if not, which) K-rings are commutator-nil; it would be particularly interesting to know the answer for the subclass of Engel rings.

The results of this section have so far applied only to certain special kinds of K-ring. We now establish a property common to all K-rings:

THEOREM 4.4. *Every K-ring is commutative modulo its Jacobson radical.*

Proof. By Theorem 3.2, we may confine ourselves to showing that, if a given division ring D is also a K-ring, then it is commutative; and, by Theorem 4.3, we may suppose that D has zero characteristic. By Theorem 4.2, it will be enough to prove that, if $k(x,y)$, $n(x,y)$ are the given index functions of D as a K-ring, then $[x, y^{n(x,y)}]=0$ for all $x, y \varepsilon D$. Suppose by way of contradiction that D contained x, y such that $[x, y^{n(x,y)}]\neq 0$. On writing $y^{n(x,y)}=z$ this becomes $[x,z]\neq 0$, while we know that $e_{k(x,y)}(x,z)=0$, so there is a positive integer j ($\leqq k(x,y)-1$) such that $e_j(x,z)\neq 0$, $e_{j+1}(x,z)=0$. Since j is positive, $b=e_{j-1}(x,z)$ exists, and, writing $c=e_j(x,z)$, we have $c=[b,z]\neq 0$, $[c,z]=0$. Since $c\neq 0$, c^{-1} exists and commutes with z (since c does); hence, defining $d=-c^{-1}b$, we find $[z,d]=c^{-1}[z,-b]=1$. Writing $f=zd$, we then have $[z,f]=z\neq 0$.

We shall obtain our contradiction by showing that the relation $[z,f]=z\neq 0$ cannot hold between elements z, f of a division K-ring D having zero characteristic. We note that

$$[z,f^q]=\{(f+1)^q-f^q\}z \qquad (q=1,2,\cdots);$$

for the case $q=1$ of this merely states that $[z,f]=z$, while, if the relation holds for a given q, then

$$\begin{aligned}[z,f^{q+1}]&=f^q[z,f]+[z,f^q]f=f^q[z,f]+\{(f+1)^q-f^q\}zf\\ &=f^q[z,f]+\{(f+1)^q-f^q\}([z,f]+fz)\\ &=(f+1)^qz+\{(f+1)^q-f^q\}fz=\{(f+1)^{q+1}-f^{q+1}\}z.\end{aligned}$$

It then follows, by another trivial induction, that

$$e_i(z,f^q)=\{(f+1)^q-f^q\}^iz \qquad (i=0,1,\cdots;q=1,2,\cdots),$$

and indeed, since $[z,f+h]=[z,f]=z$, that

$$e_i(z,(f+h)^q)=\{(f+h+1)^q-(f+h)^q\}^iz$$
$$(h,i=0,1,\cdots;q=1,2,\cdots).$$

Since D is a K-ring, we can find positive integers $k=k(z,f)$, $n=n(z,f)$ such that $e_k(z,f^n)=0$, whence (since $z\neq 0$) we deduce $(f+1)^n=f^n$. But this implies

$$(f+h)^n-f^n=0 \qquad (h=1,2,\cdots);$$

for we have just proved the case $h=1$, while the result for given h entails

$$\begin{aligned}\{(f+h+1)^n-f^n\}^kz&=\{(f+h+1)^n-(f+h)^n\}^kz\\ &=e_k(z,(f+h)^n)=e_k(z,f^n)=0.\end{aligned}$$

Now the n equations $(f+h)^n-f^n=0$ ($h=1,2,\cdots,n$), regarded as homogeneous linear relations between the n quantities $C_0{}^n, C_1{}^n f, \cdots, C_{n-1}{}^n f^{n-1}$, have (Vandermonde) determinant $\Delta = 1!2!\cdots n!$; we conclude that

$$\Delta C_0{}^n = \Delta C_1{}^n f = \cdots = \Delta C_{n-1}{}^n f^{n-1} = 0,$$

which (since D has characteristic zero) gives our sought-for contradiction. Thus in fact $[x, y^{n(x,y)}] = 0$ for all x, y, and the result follows.

In the Engel case the above proof reduces almost to triviality (and is in particular independent of Herstein's result and of characteristic considerations). From well-known properties of the Jacobson radical, we have immediately

THEOREM 4. 5. *Every π-regular K-ring is commutator-nil. Every K-ring with minimal condition on its left (or right) ideals has nilpotent commutator ideal.*

5. Engel rings. The results of Section 4 can be extended either in the direction of proving something stronger than the commutator-nil property for suitably selected special K-rings, or by showing that the commutator-nil property (or commutativity modulo J) in fact holds under hypotheses less stringent than those made above. We shall obtain a few such extensions in this section and the next.

In view of the result of Jennings quoted shortly before Theorem 4. 2 above, perhaps the most conspicuous problem under the first heading is to determine for which K-rings R an integer h exists such that $[[\cdots[[x_1, x_2], x_3], \cdots], x_h] = 0$ for all choices of $x_1, \cdots, x_h$ in R (conversely, of course the existence of such an h implies that R is a K-ring, and indeed an Engel ring with $k = h - 1$ independent of x and y). Let $[R]$ denote the Lie ring associated with R, i. e. having the same additive group as R, but with the multiplication $[x, y]$. Then our problem may be restated as follows: when is the K-ring condition on R equivalent to the nilpotence of $[R]$ (as a Lie ring)?

It is known that every finite-dimensional Engel algebra R has $[R]$ nilpotent. Also, even in the wider class of arbitrary Lie rings L (i. e. not only those which come from associative rings in the way just described), it is known that the Engel condition $e_k(x, y) = 0$, with index function $k = k(x, y)$, implies the nilpotence of L whenever either (1) L is soluble and finitely generated, or (2) L is soluble and of zero characteristic and $k(x, y)$ is independent of both x and y, or (3) L satisfies the maximal condition on (Lie) subrings and $k(x, y)$ is independent of x. These three results are respectively due to Gruenberg [8; Theorem 1′], Higgins [11; Theorem 1] and Zorn [19];

Higgins also showed [Theorem 4] that zero characteristic and the Engel condition with $k(x, y) \leqq 4$ together always ensure solubility. We now ask how far these provisos may be dispensed with when L is the Lie ring $[R]$ of some associative ring R.

We certainly cannot hope to prove that every associative Engel ring R has $[R]$ nilpotent: for Cohn [4] has constructed, for each prime p, an associative ring of characteristic p in which $e_{p+1}(x, y)$ vanishes identically but which contains an infinite sequence of elements x_j such that $[[\cdots[x_1, x_2], \cdots], x_j] \neq 0$ $(j = 2, 3, \cdots)$. However, we do at any rate have

THEOREM 5.1. *Let R be any given associative Engel ring, with index function $k(x, y)$ independent of both x and y, and suppose further that R has zero characteristic and satisfies the minimal or maximal condition on left (or right) ideals; then the Lie ring $[R]$ is nilpotent.*

Proof. Every nil two-sided ideal of any associative ring with minimal or maximal condition on left (or right) ideals is nilpotent; this is well known in the minimal case, and was proved by Levitzki [17] for the maximal case. Hence, by Theorem 4.1, R_1 is nilpotent, and so the Lie ring $[R]$ is clearly soluble. The nilpotence of $[R]$ now follows from Higgins' result (2).

We could, if we wished, strengthen the conclusion of Theorem 5.1 by asserting that R is even of finite class, i.e. that the lower central series $R = R_0 \geqq R_1 \geqq R_2 \geqq \cdots$ as defined by Jennings [13] terminates in zero; this follows from Jennings' Theorem 6.5. Also, by Gruenberg's result (1) and [7; Theorem 2], we see that every finitely generated associative Engel ring R with R_1 nilpotent has finite class. Now, if a given Engel ring R either has $k = k(y)$ or has prime characteristic or is π-regular, then we already know that R_1 is nil; hence, if R also satisfies the maximal condition on left (or right) ideals, then R_1 must be nilpotent, and our previous remark shows that R must have finite class locally. Again, by Theorem 4.5, every Engel ring with minimal condition on left (or right) ideals has finite class locally.

Having pointed out the most obvious consequences for Engel rings of our theorems of Section 4, we now round off this section with an account of some related results of Gruenberg (which have not previously appeared in print). Given any Lie ring L, any subset X of L and any non-negative integer r, let $X_{(r)}$ denote the set of all repeated commutators $[[\cdots[x_0, x_1], \cdots], x_r]$ with every $x_j \varepsilon X$.

THEOREM 5.2. *Let R be any finitely generated Engel ring, and X any given set of generators of R (as an associative ring). Then, if the Lie subring $[X]$ generated by X in $[R]$ is soluble, R must have finite class.*

Proof. We may clearly suppose that X is itself a finite set. Then $[X]$ is a finitely generated soluble Lie subring of an Engel ring, and is accordingly nilpotent by [8; Theorem 1′]: in particular, we can find a non-negative integer i such that $X_{(i)}$ contains only the zero element of R. If $i = 0$ or 1 then of course $R_1 = 0$, while for $i \geqq 2$ [7; Theorem 2] gives

$$R_{(i-1)j} \leqq S_{(i)} + S_{(1)}{}^{j} \qquad (j = 1, 2, \cdots),$$

where $S_{(r)}$ denotes the two-sided ideal of R generated by $X_{(r)}$. Hence it is enough to prove that, for any $i \geqq 2$, $S_{(1)}{}^{j} \leqq S_{(i)}$ for sufficiently large j; by an obvious induction, we need only show that, for each positive integer r, there is a positive integer t such that $S_{(r)}{}^{t} \leqq S_{(r+1)}$.

To prove this, we observe first that, in any Engel ring, every central commutator $[a, b]$ is nilpotent. Indeed, if $[[a, b], ba] = 0$, then we easily verify that $e_j(a, ba) = [a, b]^j a$ $(j = 1, 2, \cdots)$, whence $[a, b]^{k(a, ba)} a = 0$, and similarly $[b, a]^{k(b, ab)} b = 0$; letting $k = \max\{k(a, ba), k(b, ab)\}$, we at once conclude that

$$[a, b]^{k+1} = [a, b]^k a \cdot b - (-1)^k \cdot [b, a]^k b \cdot a = 0.$$

Now define $R^* = R/S_{(r+1)}$ and, for any subset T of R, let T^* denote the image of T under the natural homomorphism of R onto R^*. Since X generates R, clearly $(X_{(r)})^* = (X^*)_{(r)}$ lies in the centre of R^*, and so, by applying the remark of the previous paragraph to the Engel ring R^*, every element of $(X^*)_{(r)}$ is nilpotent and central in R^*; consequently the ideal generated in R^* by the finite set $(X^*)_{(r)}$ is nilpotent, i.e. $S_{(r)}$ is nilpotent modulo $S_{(r+1)}$, as required.

Theorem 5.2 generalizes Jennings' result [14; Theorem 3] that, if a given finitely generated ring R has $[R]$ nilpotent, then R is of finite class. It also leads at once to

THEOREM 5.3. *If, in a given associative ring R, every finite subset X generates a nilpotent Lie ring $[X]$, then every finitely generated associative subring of R has finite class.*

To see this, we have only to apply Theorem 5.2 to the finitely generated subrings of R (noticing that the hypothesis of local Lie nilpotence always implies the Engel condition).

THEOREM 5.4. *If an Engel ring R has its index function $k(x, y)$ independent of both x and y, then R is locally of finite class.*

Proof. After Theorem 5.2, we have only to show that the Lie ring $[X]$ generated by any given finite subset X of R is soluble.

By Theorem 4.1, we know that R_1 is nil, while also, by our hypothesis on R, R_1 satisfies a polynomial identity $e_k(x,y)=0$ for some fixed k; hence, by Kaplansky [16; Theorem 5], R_1 is locally nilpotent (in the associative sense). Since X is finite we know, by [8; Theorem 2′], that the derived ring $[X]^2$ of $[X]$ is finitely generated as a Lie ring, say by the finite set Y. Then Y, being a finite subset of the locally nilpotent ring R_1, generates a nilpotent associative subring of R_1; and, since this subring of course contains $[X]^2$, clearly the Lie ring $[X]$ is soluble, as required.

We remark that the example of Baer [3] already mentioned in Section 2 shows that not every Engel ring with zero characteristic and $k=k(y)$ is locally of finite class. It remains open whether every Engel ring R with zero characteristic and $k(x,y)$ independent of both x and y has $[R]$ nilpotent (or, equivalently, soluble).

6. **Weak *K*-rings.**

To extend our K-ring theorems of Section 4 to more general types of ring, we turn first to generalizations within the class of m-rings; since every division ring is an m-ring for all m, we exclude some of them by requiring in the defining property that $\alpha=0$, $a=0$, i.e. we confine attention to those m-rings ($m=1,2,\cdots$) to each choice of x,y in which there correspond integers k,l,n such that $1\leqq l\leqq m$, $x^{m-l}e_k(-x^l,y^n)=0$. We shall in fact consider only the case $l=m$, but it turns out that we can usefully generalize this specialization of the m-rings by allowing m to depend on x,y: formally, we call a given ring R a *weak K-ring* whenever, to each choice of x,y in R, there correspond positive integers $k=k(x,y)$, $m=m(x,y)$ and $n=n(x,y)$ such that

$$e_k(x^m,y^n)=0.$$

Obviously every K-ring is a weak K-ring. We note also that certain weak K-rings which do not at first sight seem to be K-rings nevertheless turn out to be on closer inspection:

THEOREM 6.1. *Let R be a weak K-ring of characteristic zero, whose index functions $k(x,y)$, $m(x,y)$, $n(x,y)$ are either* (a) *all independent of x, or* (b) *all independent of y. Then R is a K-ring with respectively either* (a) $e_{k(x,y)}(x,y^{m(x,y)n(x,y)})=0$ *or* (b) $e_{k(y,x)}(x,y^{m(y,x)n(y,x)})=0$ *for all x,y; in particular, by Theorem 4.1, R is commutator-nil.*

Proof. (a) Take any y in R (fixed throughout), and write for brevity

$$k(x,y)=k,\ m(x,y)=m,\ n(x,y)=n,\ y^n=z,$$

so that, for any $x,u\,\varepsilon\,R$, $e_k((x+hu)^m,z)=0$ $(h=0,1,\cdots,m-1)$. Since

R has characteristic zero we deduce by the familiar "linearization argument" (cf. for example Higgins [11]) that

$$e_k(xu^{m-1}+uxu^{m-2}+\cdots+u^{m-1}x,z)=0.$$

If R has a unit element 1, then, on taking $u=1$ in this, we obtain $e_k(mx,z)=0$, i.e. $e_k(x,z)=0$, whence it is easy to see that $e_k(x,z^m)=0$; we have to show that this conclusion holds independently of the presence of a unit.

Now clearly $[xz^{m-1}+zxz^{m-2}+\cdots+z^{m-1}x,z]=[x,z^m]$, and so, taking $u=z$, we have

$$\begin{aligned}e_{k-1}([x,z^m],z)&=e_{k-1}([xz^{m-1}+\cdots+z^{m-1}x,z],z)\\&=e_k(xz^{m-1}+\cdots+z^{m-1}x,z)=0.\end{aligned}$$

Also we can easily verify (by induction) that, for any $b\,\varepsilon\,R$,

$$e_j(b,z^m)=\sum_{i_1=1}^{m}\cdots\sum_{i_j=1}^{m}z^{i_1+\cdots+i_j-j}e_j(b,z)z^{jm-i_1-\cdots-i_j}\qquad(j=1,2,\cdots);$$

in particular, $e_j(b,z^m)$ vanishes whenever $e_j(b,z)$ does. Hence, taking $j=k-1$, $b=[x,z^m]$, we find $e_k(x,z^m)=e_{k-1}([x,z^m],z^m)=0$, as required.

(b) Take any y in R (fixed throughout), and write now

$$k(y,x)=k,\ m(y,x)=m,\ n(y,x)=n,\ y^m=z,$$

so that, for any $x,u\,\varepsilon\,R$, $e_k(z,(x+hu)^n)=0\quad(h=0,1,\cdots,kn-1)$. Linearizing once more and singling out the terms linear in x, we find

$$\sum_{i=0}^{k-1}e_{k-i-1}([e_i(z,u^n),xu^{n-1}+\cdots+u^{n-1}x],u^n)=0.$$

On putting $u=z$ in this, each summand with $i\neq 0$ clearly reduces to 0; hence the summand corresponding to $i=0$ must also vanish under the substitution $u=z$, i.e. we have

$$\begin{aligned}0&=e_{k-1}([e_0(z,z^n),xz^{n-1}+\cdots+z^{n-1}x],z^n)\\&=e_{k-1}([z,xz^{n-1}+\cdots+z^{n-1}x],z^n)=e_{k-1}([z^n,x],z^n)\\&=-e_{k-1}([x,z^n],z^n)=-e_k(x,y^{mn}),\end{aligned}$$

as required.

Theorem 6.1 (b), together with Lemma 4.2 and Theorem 4.3, enables us to make the following addition to our collection of K-ring results:

THEOREM 6.2. *Every K-ring whose index functions k,n depend only on x is commutator-nil.*

To show that every weak K-ring with k, m, n depending only on x (or only on y) is commutator-nil, it would, by Lemma 4.2, be enough to deal with the cases of zero and prime characteristic. Theorem 6.1 settles the first case, while in the second case the argument of Lemma 4.1 (ii) shows that we may take $k=1$; thus it would even suffice to prove that every ring with prime characteristic and $[x^{m(x)}, y^{n(x)}]=0$ is necessarily commutator-nil. But the writer has not yet been able to settle this critical case.

Our next two results extend Theorem 2.2, and the main part of the argument of Theorem 4.4, to all weak K-rings.

THEOREM 6.3. *Every primitive weak K-ring is a division ring.*

Proof. This is trivial (and still along essentially the same lines as that of Theorem 2.2). We may take our ring to be a dense set M of (right) linear transformations of a vector space V over a division ring D, and have only to show that V cannot contain two D-independent elements v_0, v_1. Now, if such v_0, v_1 existed, we could, since M is dense on V, find x, y in M such that $v_0x=v_1$, $v_1x=v_1$, $v_0y=0$, $v_1y=v_1$; but then, for any k, m, n,

$$v_0e_k(x^m, y^n)=v_0x^my^{kn}=v_1y^{kn}=v_1\neq 0$$

(since v_0, v_1 are D-independent), which is impossible if M is a weak K-ring.

THEOREM 6.4. *If a division ring D has the weak K-property, i.e. if there correspond to each pair x, y of elements of D integers $k(x,y)$, $m(x,y)$, $n(x,y)$ such that $e_{k(x,y)}(x^{m(x,y)}, y^{n(x,y)})=0$, then in fact $[x^{m(x,y)}, y^{n(x,y)}]=0$ for all x, y, i.e. we may always replace $k(x,y)$ by 1 while keeping $m(x,y)$ and $n(x,y)$ as given.*

Proof. Since D is a division ring, it has characteristic either a prime or zero; we consider these two cases separately.

If D has prime characteristic p, then, given any $x, y\in D$, and writing (for this half of the proof only) $m(x,y)=m$, $n(x,y)=n$, we can find a positive integer $r=r(x,y)$ such that $p^r\geqq k(x,y)$; then, by the argument of Lemma 4.1 (ii), we have

$$[x^m, (y^n)^{p^r}]=e_{p^r}(x^m, y^n)=0.$$

If $x\neq 0$, then x^{-m} exists, and

$$\begin{aligned}0=x^{-m}[x^m, (y^n)^{p^r}] &= (y^n)^{p^r}-x^{-m}(y^n)^{p^r}x^m\\ &=(y^n)^{p^r}-(x^{-m}y^nx^m)^{p^r}=(y^n-x^{-m}y^nx^m)^{p^r};\end{aligned}$$

hence, since D is a division ring,

$$0 = y^n - x^{-m}y^nx^m = x^{-m}[x^m, y^n],$$

i. e. $[x^{m(x,y)}, y^{n(x,y)}] = 0$.

Suppose on the other hand that D has zero characteristic. If x, y existed in D with $[x^{m(x,y)}, y^{n(x,y)}] \neq 0$, then, writing $y^{n(x,y)} = z$, we could (as in the proof of Theorem 4. 4) find a positive integer j such that $c = e_j(x^{m(x,y)}, z) \neq 0$, $[c, z] = 0$; on writing $d = -c^{-1}e_{j-1}(x^{m(x,y)}, z)$, it would follow from this that $[z, d] = 1$. Again, writing $zd = f$, we proceed to deduce a contradiction from $[z, f] = z \neq 0$. We now start from the relations

$$[z^s, f^q] = \{(f+s)^q - f^q\}z^s \qquad (q, s = 1, 2, \cdots),$$

which can be verified by induction on q as before (though the case $q = 1$ is no longer vacuous); it follows at once that

$$e_i(z^s, (f+hs)^q) = \{(f+(h+1)s)^q - (f+hs)^q\}^i z^s \qquad (h, i = 0, 1, \cdots; q, s = 1, 2, \cdots).$$

Since D is a K-ring, $e_k(z^m, f^n) = 0$ for suitable positive integers k, m, n, and so

$$(f+hm)^n - f^n = 0 \qquad (h = 1, 2, \cdots, n),$$

whence, with $\Delta = 1!\,2!\cdots n!$ as before,

$$\Delta C_0{}^n m^n = \Delta C_1{}^n m^{n-1} f = \cdots = \Delta C_{n-1}{}^n m f^{n-1} = 0,$$

our desired contradiction.

In view of Theorems 6. 3 and 6. 4, to prove *every* weak K-ring commutative modulo its Jacobson radical, it would suffice to show that every division ring D subject to $[x^{m(x,y)}, y^{n(x,y)}] = 0$ (for all $x, y \in D$) is commutative. This is yet another unsolved problem; indeed, it is not even known whether commutativity holds when D has prime characteristic and m, n are either (a) both independent of x, or equivalently (b) both independent of y (which would at any rate imply that every weak K-ring R with m, n satisfying (a) or (b) has R/J commutative). However, by Theorem 4. 2, we can at least assert

THEOREM 6. 5. *Let R be any ring over which are defined integral-valued functions $k(x, y)$, $m(x, y)$ such that $e_{k(x,y)}(x^{m(x,y)}, y) = 0$ for all $x, y \in R$. Then R is commutative modulo its Jacobson radical.*

TRINITY COLLEGE, CAMBRIDGE.

REFERENCES.

[1] A. S. Amitsur, *Journal of the London Mathematical Society*, to appear.

[2] R. F. Arens and I. Kaplansky, "Topological representation of algebras," *Transactions of the American Mathematical Society*, vol. 63 (1948), pp. 457-481.

[3] R. Baer, "Radical ideals," *American Journal of Mathematics*, vol. 65 (1943), pp. 537-568.

[4] P. M. Cohn, "A non-nilpotent Lie ring satisfying the Engel condition and a non-nilpotent Engel group," *Proceedings of the Cambridge Philosophical Society*, vol. 51 (1955), pp. 401-405.

[5] M. P. Drazin, "Rings with central idempotent or nilpotent elements," to appear.

[6] ———, "A generalization of polynomial identities in rings," to appear.

[7] M. P. Drazin and K. W. Gruenberg, "Commutators in associative rings," *Proceedings of the Cambridge Philosophical Society*, vol. 49 (1953), pp. 590-594.

[8] K. W. Gruenberg, "Two theorems on Engel groups," *Proceedings of the Cambridge Philosophical Society*, vol. 49 (1953), pp. 377-380.

[9] I. N. Herstein, "A thorem on rings," *Canadian Journal of Mathematics*, vol. 5 (1953), 238-241.

[10] ———, "The structure of a certain class of rings," *American Journal of Mathematics*, vol. 75 (1953), pp. 864-871.

[11] P. J. Higgins, "Lie rings satisfying the Engel condition," *Proceedings of the Cambridge Philosophical Society*, vol. 50 (1954), pp. 8-15.

[12] N. Jacobson, *The Theory of Rings* (American Mathematical Society, 1943).

[13] S. A. Jennings, "Central chains of ideals in an associative ring," *Duke Mathematical Journal*, vol. 9 (1942), pp. 341-355.

[14] ———, "On rings whose associated Lie rings are nilpotent," *Bulletin of the American Mathematical Society*, vol. 53 (1947), pp. 593-597.

[15] I. Kaplansky, "A theorem on division rings," *Canadian Journal of Mathematics*, vol. 3 (1951), pp. 290-292.

[16] ———, "Rings with a polynomial identity," *Bulletin of the American Mathematical Society*, vol. 54 (1948), pp. 575-580.

[17] J. Levitzki, "Solution of a problem of G. Koethe," *American Journal of Mathematics*, vol. 67 (1945), pp. 437-442.

[18] M. Nagata, "On the nilpotency of nil-algebras," *Journal of the Mathematical Society of Japan*, vol. 4 (1952), pp. 296-301.

[19] M. Zorn, "On a theorem of Engel," *Bulletin of the American Mathematical Society*, vol. 43 (1937), pp. 401-404.

A CHARACTERIZATION OF PLANE CONTINUOUS CURVES WHICH ARE FILLED UP BY CONTINUOUS COLLECTIONS OF CONTINUOUS CURVES.*

By Mary-Elizabeth Hamstrom.

1. Introduction and notation. Let M denote a compact continuous curve in the plane. The purpose of this paper is to establish necessary and sufficient conditions on M that it should be the sum of the elements of a continuous collection G of mutually exclusive nondegenerate continuous curves. It is well known that if such a collection does fill up M then G is a continuous curve with respect to its elements and it is proved in [5] that G is an arc or a simple closed curve. Theorem 3 of [3] states that there is a continuous collection of arcs which fills up M and is an arc with respect to its elements if and only if M is a closed disc (a simple closed curve plus its interior) and Theorem II of [1] states that there is a continuous collection of simple closed curves filling up M if and only if M is an annulus. (In this connection, see also [2].)

This paper will be restricted to the consideration of points in the plane, E. Throughout this paper the letter M will be used to denote a compact continuous curve in E, $B(M)$ will denote the boundary of M and $U(M)$ will denote the collection of all boundaries of complementary domains of M. The notation $T(M)$ will denote the collection of all nondegenerate subcontinua x of M such that x is maximal with respect to the property of being a continuum which is the limiting set of an infinite sequence of distinct elements of $U(M)$. The notation $S(M)$ will denote the collection consisting of the elements of $U(M)+T(M)$ and all points[1] of $B(M)-[U^*(M)+T^*(M)]$.

Definition 1. If S is an upper semicontinuous collection of continua in the plane and if S, with respect to its elements, is an arc with endelements a and b then an arc h which is a non-endelement of S is said to be a K-arc or K-curve of S provided there exist two simple closed curves, J_1 and J_2,

---

* Received May 12, 1954; revised June 23, 1955. Presented to the Society, December 30, 1953.

[1] If K is a collection of point sets, K^* will denote the sum of the elements of K.

with mutually exclusive interiors, I_1 and I_2, such that $J_1 \cdot J_2$ is an arc which contains and is separated by h, $J_1 \cdot J_2 \cdot S^* = h$, I_1 does not intersect any element of the interval[2] ah of elements of S and I_2 does not intersect any element of the interval hb. If h is a or b then h is a K-curve provided there is a simple closed curve containing h but intersecting no other element of S. A simple closed curve which is an element of S is said to be a K-curve of S provided it either separates two points of S^* in E or is an endelement of S. A point which is an element of S is said to be a K-curve of S provided it is not the sequential limiting set of any sequence of simple closed curves in S.

Definition 2. If V is a closed subcollection of arcs of S then if n is a natural number $V^{(n)}$ will denote the collection of all elements g of V such that if ϵ is a positive number and U is an open subcollection of S containing g then U contains an element of V which has a subarc the distance between whose endpoints is less than ϵ but whose diameter is not less than $1/n$.

Definition 3. A collection S of mutually exclusive point sets is said to be imbedded in a collection W of mutually exclusive point sets provided that each element of S is a subset of some element of W and each element of W contains one and only one element of S. (These definitions were used in [6].)

2. Statement of the principal theorem, some examples and some properties of *M*.

THEOREM 2. *The compact plane continuous curve M is the sum of the elements of a continuous collection of mutually exclusive nondegenerate continuous curves which is an arc with respect to its elements if and only if every boundary of a complementary domain of M is a simple closed curve and there is an upper semicontinuous collection S such that* (1) *S is an arc with respect to its elements,* (2) *$B(M)$ is a subset of S^*,* (3) *every non-endelement of S is either an element of $T(M)$, a point of $M - B(M)$, a point of $B(M) - [U^*(M) + T^*(M)]$ or a point or pair of points of some element of $U(M)$,* (4) *each endelement of S is a nondegenerate continuous curve which does not separate M and is a continuum of condensation*[3] *of M,* (5) *each non-endelement of S which is an element of $T(M)$ is an arc or*

---

[2] If H is an upper semicontinuous collection of point sets which is an arc with respect to its elements and x and y are elements of H then the interval xy of elements of H is the collection consisting of x, y, and all elements of H between x and y. The segment xy is the interval xy minus the elements x and y.

[3] A nondegenerate subcontinuum K of M is a continuum of condensation of M provided each point of K is a limit point of $M - K$.

simple closed curve which is a K-curve of S and does not intersect any element of $U(M)$ in more than one point and (6) *if H is a closed subcollection of arcs which are elements of S then for no natural number n is it true that $H^{(n)} = H$.*

Theorem 2 gives a condition on the boundary of M. Example 4 below, although fairly elementary, shows that condition (4) in the statement of Theorem 2 permits the boundary of M to have some rather complicated components and so long as these components can be made subsets of the endelements of S the conditions (1) through (6) can be satisfied. This prevents the formulation of a less complicated characterization.

Suppose V is an upper semicontinuous collection of mutually exclusive arcs, simple closed curves and points which fills up a compact plane continuum and is an arc with respect to its elements. In [6] (Theorem 2) I proved that there exists (in the plane) a continuous collection W of simple closed curves such that V is imbedded in W if and only if (1) each element of V is a K-curve of V, (2) if H is a closed subcollection of arcs of V then for no natural number n is it true that $H^{(n)} = H$ and (3) of each pair of simple closed curves which are elements of V one is interior to the other. This theorem will be applied in a certain way to the elements of S in order to show that the conditions of the statement of Theorem 2 are sufficient.

Example 1. Suppose M consists of two concentric circles C_1 and C_2 plus the points between them. Let t_1 and t_2 be mutually exclusive arcs lying in $M - (C_1 + C_2)$. Let P_1 and Q_1 and P_2 and Q_2 be pairs of diametrically opposite points of C_1 and C_2 respectively. Let s_1 and s_2 be mutually exclusive arcs each lying except for its endpoints in $M - (C_1 + C_2 + t_1 + t_2)$ where one endpoint of s_i is P_i, the other a point of t_i $(i = 1, 2)$. Let s be an arc lying except for its endpoints Q_1 and Q_2 in $M - (C_1 + C_2 + t_1 + t_2 + s_1 + s_2)$. Then if S is the collection consisting of t_1, t_2, the points of $s_1 + s_2 + s$ not in $t_1 + t_2$, and the pairs of points which are the endpoints of the chords of C_i perpendicular to the diameter P_iQ_i $(i = 1, 2)$, the conditions (1) through (6) of Theorem 2 are satisfied. If C_2 is interior to C_1 then the collection G may consist of (1) t_1 and t_2, (2) a simple closed curve g_1 having $C_2 + s_2 + t_2 + s$ in its exterior and $t_1 + s_1 - P_1$ in its interior and such that $g_1 \cdot C_1 = g_1 \cdot s_1 = P_1$, (3) a simple closed curve g_2 having $g_1 + t_1 + s_1$ in its exterior and $C_2 + s_2 + t_2 + s - Q_1$ in its interior and such that $g_2 \cdot C_1 = Q_1$, (4) a simple closed curve g_3 having $g_2 + s - Q_2$ in its exterior and $t_2 + s_2 + C_2 - Q_2$ in its interior and such that $g_3 \cdot C_2 = Q_2$, (5) a simple closed curve g_4 having $g_3 + C_2 - P_2$ in its exterior and $t_2 + s_2 - P_2$ in its interior and such that $g_4 \cdot C_2 = P_2$, (6) appropriate simple closed curves each lying interior to g_1 or g_4, each having t_1 or t_2 in its interior and each having only one point in common with s_1 or s_2, (7) appropriate arcs each either lying exterior to both g_1 and g_2 and having for endpoints the endpoints of a chord of C_1 perpendicular to P_1Q_1 or lying interior to g_3 and exterior to g_4 and having for endpoints the endpoints of a chord of C_2 perpendicular to P_2Q_2, and (8) appropriate simple closed curves lying interior to g_2 and exterior to g_3 and having only one point in common with s.

In the more general cases the construction of G will be patterned after the above construction, which is a natural way to proceed.

Example 2. Let M' be a simple closed curve J plus its interior I, let K denote a contracting sequence of mutually exclusive closed discs such that K^* is a subset of I which is dense in I and let M denote M' minus the sum of the interiors of the elements of K. Clearly there is no collection satisfying the conditions (1) through (6) in the statement of Theorem 2.

Example 3. Let t be an arc lying in the interior I of a closed disc M', let α be a contracting sequence of mutually exclusive closed discs such that t is the limiting set of α and α^* lies in $I - t$ and let M denote M' minus the sum of the interiors of the elements of α. If we choose t as one of the elements of $T(M)$ it is seen that in this case a collection S satisfying conditions (1) through (6) does exist.

Example 4. Let J be a simple closed curve lying in the interior of a closed disc M' and let J_1, J_2, and J_3 be mutually exterior simple closed curves each lying in the interior of M', each having one and only one point in common with J and each having its interior lying in the exterior of J. Let M denote M' minus the sum of the interiors of J, J_1, J_2, and J_3. It is seen that if $J_1 + J_2 + J_3 + J$ is an endelement of S then a collection satisfying the conditions of Theorem 2 does exist. There is a continuous collection of continuous curves filling up M, one endelement of which is $J + J_1 + J_2 + J_3$, the other endelement is the outer boundary of M and the non-endelements are simple closed curves whose interiors contain $J + J_1 + J_2 + J_3$.

Example 5. It has been suggested by a referee that condition (6) of Theorem 2 can be omitted. The author even announced at the April, 1955 meeting of the Society in Brooklyn, N. Y. that if V is an upper semicontinuous collection of arcs and points such that (1) V is an arc with respect to its elements and fills up a compact plane continuum and (2) each nondegenerate element of V is a K-curve of V and a continuum of condensation of V^* then if H is a closed subcollection of arcs of V then for no natural number n is it true that $H^{(n)} = H$. If this were a true statement condition (6) would follow immediately from conditions (1) through (5) and the above mentioned Theorem 2 of [6]. However, the following example shows that this statement is not true.

Let T be a Cantor set in the interval $(0,1)$ of the X axis. If x is a point of $(0, 1)$ let T_x be a vertical straight line interval from the point $(x, 0)$ to the point $(x, 1)$. Let $I_1, I_2, \cdots$ be the components (which are segments) of $(0,1) - T$, let $x_1, x_2, \cdots$ and $z_1, z_2, \cdots$ be the right and left hand endpoints of these segments and let $y_1, y_2, \cdots$ be their midpoints. For each j let S_{x_j} be an arc which is the sum of T_{x_j}, T_{y_j} and the straight line interval from the point $(y_j, 1)$ to the point $(x_j, 0)$. Let H be the collection consisting of S_{x_j} for each j and T_x for each x in $T - \Sigma x_j$. For each j let G_j be an upper semicontinuous collection which is an arc with respect to its elements such that (1) each non-endelement of G_j is a point, (2) the end-elements of G_j are S_{x_j} and T_{z_j}, (3) each point of $S_{x_j} - T_{x_j}$ is a limit point of $G_j^* - S_{x_j}$ but no point of $T_{x_j} - x_j$ is such a limit point and (4) every point of G_j^* lies on or between the lines $y = 0$ and $y = 1 + 1/i$ and on or between the lines $x = z_j$ and $x = x_j$. The collection consisting of the elements of ΣG_i and H is an upper semicontinuous collection V of arcs and points, it is an arc with respect to its elements and each of its nondegenerate elements is a K-curve of V and a continuum of condensation of V^*. However, $H^{(2)} = H$.

If J is a simple closed curve whose interior contains V^* and appropriate intervals of points of G_j are replaced by appropriate closed discs whose interiors are subsequently removed from the interior of J, we obtain an example of a continuous curve satisfying all the conditions of Theorem 2 except (6). A study of the construction of V in the above paragraph should convince the reader that this second example can be easily obtained.

Before proceeding to the proof of Theorem 2 I shall establish some lemmas. In what follows, the letter G will denote a continuous collection of nondegenerate continuous curves which fills up M and is an arc with respect to its elements.

LEMMA 1. *Every complementary domain of M is bounded by a simple closed curve and no two such boundaries have more than one point in common.*

Proof. It follows easily that M has no cut point. Theorem 54, Chapter IV of [8] states that if K is the boundary of a complementary domain D of a compact continuum N and A is a cut point of K then A is a cut point of N.

Consequently, no boundary of a complementary domain of M has a cut point and since M is locally connected every such boundary curve is a simple closed curve. Clearly no two such boundary curves have more than one point in common.

COROLLARY. *If the non-endelement g of G intersects the element J of $U(M)$, (the collection of all boundary simple closed curves of M), then $g \cdot J$ does not contain an arc.*

Proof. This corollary is a rather easy consequence of the fact that the elements of G and the closures of the components of $M - g$ are locally connected. A detailed proof may be found in Lemma 1 of [4].

LEMMA 2. *Each non-endelement of G is either an arc or a simple closed curve.*

Proof. Let g be a non-endelement of G and let U be an arc of elements of G one of whose endelements is g. The continuous collection U fills up the continuous curve U^*, hence, by Lemma 1, every boundary of a complementary domain of U^* is a simple closed curve. But since g is a limit element of $G - U$, g is a subset of such a boundary simple closed curve and is thus either an arc or a simple closed curve.

The following theorem follows directly from Theorem 8 of [4] but since I use the result quite frequently I shall prove it here in a modified form. It shows how the elements of G are related to each other and to the boundary simple closed curves of M.

THEOREM 1. *Let J be an element of $U(M)$ that is not a subset of either endelement of G and let K be the collection of all elements of G which intersect J. Then (1) each non-endelement of K is an arc g such that $g \cdot J$ is the sum of the endpoints of g, (2) no non-endelement of K intersects any element of $U(M)$ other than J, (3) each endelement h of K either is an endelement of G and $h \cdot J$ is connected or is a simple closed curve which is a non-endelement of G, does not intersect three elements of $U(M)$ and has only one point in common with J.*

Proof. The collection K is clearly an arc with respect to its elements. If the common part of J and the endelement h of K separates J then h separates two elements of K in M and consequently in K^*. This is impossible, so $h \cdot J$ is connected and if h is a non-endelement of G (and hence is an arc or a simple closed curve) it follows from the corollary to Lemma 1 that $h \cdot J$ is degenerate.

If g is a non-endelement of K, $g \cdot J$ separates J and is therefore not connected and, by the corollary to Lemma 1, $g \cdot J$ is totally disconnected. By Lemma 2, g is an arc or a simple closed curve. If $g \cdot J$ contains three points or if g is a simple closed curve, $M - g$ is the sum of three mutually separated point sets, which is not true. Hence g is an arc and it follows from the continuity of the collection G that the two points of $g \cdot J$ are the endpoints of g.

Suppose the above mentioned non-endelement g of K intersects an element J' of $U(M)$ different from J. Let K' denote the collection of all elements of G which intersect J'. (By the corollary to Lemma 1, K' is a nondegenerate collection.) Since g belongs to both K and K' and is a non-endelement of K there is an element g' of G which is a non-endelement of both K and K'. Since J and J' can not both contain the endpoints of g' a contradiction follows.

If an arc k lies in and separates M then k contains two points of some element of $U(M)$. For, if not, then if $M - k$ is the sum of the two mutually separated point sets V_1 and V_2 every element of $U(M)$ lies either in $V_1 + k$ or $V_2 + k$. Consequently $E - k$ is the sum of two mutually separated point sets U_1 and U_2 where U_i $(i = 1, 2)$ is V_i plus the sum of the complementary domains of M whose boundaries lie in $V_i + k$. Since no arc separates E this is impossible.

Consequently, if the endelement h of K is an arc and a non-endelement of G, h contain two points of an element J' of $U(M)$ and is a non-endelement of the collection of elements of G which intersect J'. But no such non-endelement intersects two boundary simple closed curves. Therefore h is a simple closed curve.

If the endelement h of K is a non-endelement of G (therefore a simple closed curve) and intersects three elements, J, J_1, and J_2 of $U(M)$, two of these elements—say J_1 and J_2—have the property that $J_1 - J_1 \cdot h$ and $J_2 - J_2 \cdot h$ lie in the same complementary domain of h. Since G is a continuous collection, some arc which is a non-endelement of G intersects both J_1 and J_2, which we have seen to be impossible. This completes the proof of Theorem 1.

It is to be noted that the interior of each simple closed curve which is a non-endelement of G contains an endelement of G. Thus the collection of non-endelements does not contain three mutually exterior simple closed curves. Also, if h and h' are non-endelements of G and are also the endelements of the collection K considered above, then the interiors of h, h' and J are not mutually exclusive. If this is so then each element of $G - K$ is either in the interior of h or in the interior of h' and the non-endelements of K are exterior to both h and h'. However, the outer boundary of M is not interior to h or h' and is not J. Hence a non-endelement of K intersects both J and the outer boundary of M, which Theorem 1 has shown to be impossible. Hence the interiors of h, h' and J are not mutually exclusive. In particular, if J is not the outer boundary of M then one of h and h' is interior to the other.

LEMMA 3. *If the limiting set K of a sequence α of distinct elements of $U(M)$ is a nondegenerate continuum then K is a subset of some element of G.*[4]

Proof. Since M is locally connected and no two of its boundary simple closed curves have more than one point in common, every sequential limiting set of a subsequence of α is degenerate. Let G' denote the collection of all elements of G which intersect K. Let J be an element of α lying in G'^* but intersecting neither endelement of G and let V denote the collection of non-endelements of the collection of all elements of G intersecting J. The set V^* is open in M, but since no element of V intersects any element of $U(M)$ other than J, V^* does not intersect any element of $U(M)$ other than J. Since K intersects V^* and is the limiting set of α this is impossible. Hence K is a subset of some element of G.

LEMMA 4. *If the element J of $U(M)$ does not intersect either endelement of G then J does not contain three limit points of $B(M)-J$ (where $B(M)$ denotes the boundary of M).*[5]

Proof. Let H denote the collection of all elements of G which intersect J. It follows from the corollary to Lemma 1 that H is not degenerate. The sum, V, of the non-endelements of H is an open subset of M and it follows from Theorem 1 that V does not intersect any element of $U(M)$ different from J. Hence V does not intersect $B(M)-J$. Thus, if J contains limit points of $B(M)-J$ these points lie in the endelements of H. But by Theorem 1, the sum of the endelements of H has only two points in common with J. Thus Lemma 4 is proved.

LEMMA 5. *If a non-endelement u of G contains an element of $T(M)$, u is a simple closed curve.* (*See the definition of $T(M)$ in Section* 1.)

Proof. It follows from Lemma 2 that u is either an arc or a simple closed curve. If u is an arc the argument in the fourth paragraph of the proof of Theorem 1 proves that there is a boundary simple closed curve J of M such that $u \cdot J$ is not connected. Thus, by Theorem 1, u is a non-endelement of the collection H of all elements of G which intersect J. But by Theorem 1 again, V, the sum of the non-endelements of H, does not intersect any element of $U(M)$ other than J, which, since u contains the limiting set of an infinite sequence of distinct elements of $U(M)$, is impossible. Therefore u is a simple closed curve.

---

[4] This also follows from Theorem 9 of [4].

[5] This is a slight extension of Lemma 2 of [4].

It is to be noted that it follows from Lemma 3 that every element of $T(M)$ is a subset of some element of G.

LEMMA 6. *If K_1 and K_2 are continua such that* (1) *$K_1 + K_2$ is a subset of $B(M)$* (*the boundary of M*) *and does not intersect either endelment of G,* (2) *the common part of K_1 and K_2 is either a point or an element of $T(M)$ and* (3) *each of K_1 and K_2 contains an element of $U(M)$ and is the sum of some elements of $S(M)$, then $K_1 \cdot K_2$ does not contain a limit point of $B(M) - (K_1 + K_2)$.*

Proof. Suppose a point P of $K_1 \cdot K_2$ is a limit point of $B(M) - (K_1 + K_2)$ and let g denote the element of G which contains P. If P is a point of an element t of $T(M)$ then t is a subset of g, by Lemma 3. In any case, since $U^*(M)$ is a dense subset of $B(M)$, P is a limit point of the common part of $B(M) - (K_1 + K_2)$ and $U^*(M)$. Also, one of the two components of $M - g$—say L—has the property that P is a limit point of the common part of $B(M) - (K_1 + K_2)$, L and $U^*(M)$. Thus there is a sequence $J_1, J_2, \cdots$ such that P is a limit point of ΣJ_i and for each i, J_i belongs to $U(M)$ and intersects the common part of L and $B(M) - (K_1 + K_2)$.

Suppose one of the continua K_1 and K_2—say K_1—intersects L. There is a natural number n and a continuum U of elements of G such that every element of U is an arc which intersects J_n (and thus is a non-endelement of the collection of all elements of G which intersect J_n) and U^* separates some point of $K_1 \cdot L$ from $K_1 \cdot K_2$ in $\bar{L}$. It follows from the hypothesis to this lemma that some element of $U(M)$ intersects $K_1 \cdot U^*$, which, since U^* also intersects J_n is, by Theorem 1, impossible.

Therefore both K_1 and K_2 are subsets of g plus the other component, N, of $M - g$. There is an element J of $U(M)$ lying in K_1 and a subcontinuum U of elements of G such that every element of U is an arc which intersects J and U^* separates some point of $K_2 \cdot N$ from $K_1 \cdot K_2$ in $\bar{N}$. But K_2 contains an element of $U(M)$ which intersects U^*—which, again by Theorem 1, is impossible. Thus the supposition that $K_1 \cdot K_2$ contains a limit point of $B(M) - (K_1 + K_2)$ leads to a contradiction.

COROLLARY 1. *If K is a continuum which is a subset of $B(M)$ such that* (1) *the common part of K and an endelement h of G is either degenerate or an element of $T(M)$,* (2) *K contains an element of $U(M)$ which is not a subset of h and* (3) *K is the sum of some elements of $S(M)$ then h does not contain a limit point of $B(M) - (K + h)$.*

Proof. Since $M - h$ has only one component, the argument used in the proof of Lemma 6 can be used to prove that the assumption that h contains a limit point of $B(M) - (K + h)$ leads to a contradiction.

COROLLARY 2. *If two elements of $U(M)$ intersect h then at least one of them lies in h (where h is an endelement of G).*

Proof. If J is an element of $U(M)$ which intersects but is not contained in h, then h is an endelement of the (nondegenerate) collection of all elements of G which intersect J. But no non-endelement of this collection intersects any element of $U(M)$ other than J. Therefore, if an element J' of $U(M)$ other than J intersects h, J' is a subset of h.

LEMMA 7. *The collection T of all elements of G which contain elements of $T(M)$ is nowhere dense in G.*

Proof. Let V be an open subcollection of G. If an element v of V contains an element of $T(M)$ then V^* contains a boundary simple closed curve J of M. It follows from Theorem 1 and Lemma 5 that no non-endelement of the collection of all elements of V which intersect J contains any element of $T(M)$. The collection of all such non-endelements is an open subcollection of G. Thus every open subcollection of G contains an open subcollection of G which lies in $G - T$ and therefore T is nowhere dense in G.

It is to be noted that it is possible for an element of G to contain more than one element of $T(M)$.

3. Definition of a collection and proof that the conditions (1) through (6) of the statement of Theorem 2 are necessary. Let h and h' denote the endelements of G. If an element J of $U(M)$ intersects h or h'—say h—then $J \cdot h$ is connected, since h is an endelement of the collection H of all elements of G which intersect J. If an element t of $T(M)$ does not lie in (and consequently, by Lemma 3, does not intersect) $h + h'$, it follows from Lemmas 3 and 5 that t is a subset of a simple closed curve which is an element of G. Thus, by Theorem 1, if t intersects an element J of $U(M)$, $t \cdot J$ is degenerate.

If J is an element of $U(M)$ then (1) if J does not intersect $h + h'$

then J does not contain three limit points of $B(M)-J$, (2) if J intersects one of h and h'—say h—but not the other, $J-h\cdot J$ does not contain two limit points of $B(M)-J$, and (3) if J intersects both h and h', $J-J\cdot(h+h')$ does not contain any limit point of $B(M)-J$. Statement (1) follows from Lemma 4 and (2) and (3) follow from obvious modifications of the proof of Lemma 4.

If the element J of $U(M)$ intersects both h and h' let P_J and Q_J denote the continua $h\cdot J$ and $h'\cdot J$. If J intersects one of h and h'—say h—but not the other, let P_J denote the continuum $h\cdot J$ and if $J-h\cdot J$ contains a limit point of $B(M)-J$ (we have seen above that there are not two such points) denote it by Q_J; otherwise let Q_J denote any point of $J-h\cdot J$. If J does not intersect $h+h'$ and contains two limit points of $B(M)-J$ denote them by P_J and Q_J; if J contains only one limit point of $B(M)-J$ denote it by P_J and let Q_J denote any point of $J-P_J$; if J contains no limit point of $B(M)-J$, let P_J and Q_J denote any two points of J.

In any case, if J is not a subset of $h+h'$, $J-(P_J+Q_J)$ does not contain any limit points of $B(M)-J$ and, as was seen in the first paragraph of this section, if J intersects an element t of $T(M)$, $t\cdot J$ is a subset of P_J or Q_J. There is clearly an upper semicontinuous collection H_J of mutually exclusive point sets which fills up J, is an arc with respect to its elements and is such that its endelements are P_J and Q_J and each of its non-endelements is a pair of points separated in J by P_J+Q_J.

Let S' be the collection consisting of (1) h and h', (2) every element of $T(M)$ which does not intersect $h+h'$, (3) every element of H_J which does not lie in $h+h'+T^*(M)$, for all elements J of $U(M)$ and (4) the points of $B(M)-[h+h'+U^*(M)+T^*(M)]$. It follows from the construction of S' and the continuity of the collection G that S' is an upper semicontinuous collection.

LEMMA 8. *The collection S' contains no subcollection which is a simple closed curve with respect to its elements.*

Proof. Suppose K is a subcollection of S' which is a simple closed curve with respect to its elements. It follows from Lemma 6 and its corollaries that neither h nor h' is an element of K. Let J be an element of $U(M)$ which lies in K^* and let H be the collection of all elements of G which intersect J. The points P_J and Q_J do not lie in the same endelement of H, since neither endelement of H has two points in common with J. But, by Theorem 1, the sum of the non-endelements of H does not intersect any element of $U(M)$ other than J and hence does not intersect K^*-J, since

$U^*(M)$ is dense in $B(M)$ and therefore in S'^*. Therefore, since P_J and Q_J lie in different endelements of H the collection H_J of all elements of S' lying in J separates two elements of K in K, these two elements of K lying in different components of $M - H^*$. Since H_J is an arc with respect to its elements, this is impossible. Hence the supposition that a subcollection K of S' is a simple closed curve with respect to its elements leads to a contradiction.

It now follows from the lemmas, particularly Lemma 4 and Lemma 6 and its corollaries, that S' contains no subcollection that is a triod with respect to its elements, that every component of S' is, with respect to its elements, either an arc or a point and that if the nondegenerate component K of S' contains a limit element of $S' - K$, that limit element is an endelement of K.

Let W be the collection whose elements are the elements of S' and the points of $M - S'^*$. Clearly W is an upper semicontinuous collection filling up M, each nondegenerate cyclic element of W (the elements of W regarded as points) is homeomorphic to a two-sphere, every arc t of elements of W has a non-endelement which is a limit element of $W - t$ and no ordinary element or cyclic element of W separates W into three mutually separated subcollections. From the construction of W it is seen that every separating element of W is a simple closed curve and consequently an element of $T(M)$. Moreover, if (1) z is an arc of elements of S', considered as a subcollection of W, (2) s is an element of z, (3) z lies in a two-sphere of elements of W, and (4) s is a separating element of W then s is an endelement of z.

In a paper of R. L. Moore and J. R. Kline (see [7]) it is proved that, in the plane, in order for a closed, compact point set K to be a subset of an arc it is necessary and sufficient that every component of K should be a point or an arc t and no point of t, except, possibly, for the endpoints of t, is a limit point of $K - t$. This theorem clearly holds on a two-sphere and it is seen from the construction of the collection W and the remarks of the above paragraph that in W, S' satisfies the conditions required of K in the Moore-Kline Theorem so that this theorem can be applied to the space of elements of W. Let S be an upper semicontinuous subcollection of W such that S contains S' and S, with respect to its elements, is an arc with endelements h and h'. The collection S is the collection required by Theorem 2.

It follows directly from the construction of S that S satisfies conditions (1), (2), and (3). Condition (4) follows from the fact that h and h' are the endelements of the continuous collection G and condition (5) follows from Lemmas 3 and 5, Theorem 1 and the corollary to Lemma 1. That is, if the non-endelement t of S is an element of $T(M)$, it is, by Lemma 3, a subset of a non-endelement of G and hence is an arc or a simple closed curve. It follows from Lemma 5 that t is a subset of a simple closed curve which is an element of G. From this and the continuity of G it follows that t is a K-curve of S. The other requirements of (5) follow from Theorem 1 and the corollary to Lemma 1.

A special case of a theorem of J. H. Roberts ([9], Theorem 2) states that if E is a continuous collection of arcs filling up a compact metric space then there is a subcollection E' of E such that E' is dense in E and E is equicontinuous[6] at each element of E'. An argument suggested by that outlined in [9] yields a similar result in the case where E is a continuous collection of arcs and simple closed curves. Since each closed subcollection of arcs of S is imbedded in a continuous subcollection of arcs and simple closed curves of G then it follows from Roberts' result that if H is such a closed subcollection then for no natural number n is it true that $H^{(n)} = H$.

4. **Proof that the conditions are sufficient.**

Suppose that M is a compact plane continuous curve and that S is an upper semicontinuous collection of subsets of M such that M and S satisfy the conditions required of them in the statement of Theorem 2. Since the subcollection of S consisting of the elements of $T(M)$ is a nowhere dense subcollection it is seen that M is the closure of a domain.

Let h and h' denote the endelements of S. It follows from conditions (1) to (3) that if $h + h'$ intersects an element of $T(M)$ then $h + h'$ contains that element and if one of h and h'—say h—intersects an element J of $U(M)$ then $J \cdot h$ is connected. Let T denote the collection of elements of $T(M)$ which lie in $M - (h + h')$. From the remarks of the above paragraph it follows that T is a subcollection of S whose closure is totally disconnected in S. It also follows from conditions (1), (2), (3), and (5) that if an element t of T intersects an element J of $U(M)$ then $t \cdot J$ is degenerate.

If H_J denotes the collection of all elements of S which intersect the element J of $U(M)$ and J is not a subset of $h + h'$ then H_J is an arc with respect to its elements. Let k and k' denote the endelements of H_J. Clearly $k \cdot J$ and $k' \cdot J$ are connected. Each of k and k' is either h or h', an element of T or a point of J. Each non-endelement of H_J is a pair of points of J which separates $J \cdot k$ from $J \cdot k'$ in J.

---

[6] The collection E of mutually exclusive arcs and simple closed curves is said to be equicontinuous at an element h of E provided that for any $\epsilon > 0$ and any point p of h there exists a $\delta > 0$ such that if x and y are points of the same element h' of E and are each within δ of p then there is an arc in h' containing $x + y$ and of diameter less than ϵ.

Let K denote a two-sphere. There is a continuous transformation f carrying M onto K such that for each point x of K, $f^{-1}(x)$ is either (1) h or h', (2) an element of $S-(T+h+h')$, (3) a point of T^* or (4) a point of $M-[h+h'+U^*(M)]$. (I.e. We "sew" up the boundary simple closed curves of M.) Let $f(T)$ denote the collection of all point sets $f(t)$ where t is an element of T; let $f(S)$ denote the collection of all point sets $f(S)$ where s is an element of S.

Clearly the collection $f(S)$ is an upper semicontinuous collection of atriodic continuous curves and is an arc with respect to its elements. From the definition of K-curve for a simple closed curve and property (5) of the statement of Theorem 2 it follows that on K, if interiors be taken with respect to one of the endelements of $f(S)$, the elements of $f(S)$ which are simple closed curves have the property that of any two of them one is interior to the other. Therefore, from property (6) and a very slight modification[7] of the proof of the above mentioned Theorem 2 of [6], (see the second paragraph following the statement of Theorem 2), it follows that there is a continuous collection W filling up K such that $f(S)$ is imbedded in W, the endelements of W are the endelements of $f(S)$ and the non-endelements of W are simple closed curves. If G denotes the collection of all continua $f^{-1}(w)$ where w is an element of W, G is the collection required by Theorem 2, since, by condition (4) and the construction of f, G is continuous and each of its elements is a nondegenerate continuous curve. This completes the proof of Theorem 2.

5. Collections which are simple closed curves with respect to their elements.

THEOREM 3. *The compact plane continuous curve M is the sum of the elements of a continuous collection of nondegenerate continuous curves which is a simple closed curve with respect to its elements if and only if M is an annulus.*

Proof. The condition is obviously sufficient. To prove that it is necessary let G denote a continuous collection of nondegenerate continuous curves which fills up M and is a simple closed curve with respect to its elements and let g be an element of G. It follows from an argument in the proof of Lemma 2 that g is either an arc or a simple closed curve and it follows from the fact that g separates G locally but does not separate the continuous curve M that g is an arc. From Theorem 3 of [3] (stated in the first paragraph of this paper) it follows that G^* is the sum of two closed discs (simple closed curves plus their interiors) whose common part is the sum of two mutually exclusive arcs. Thus M is an annulus.

---

[7] Theorem 2 of [6] is proved for point sets in the plane, but the proof is easily carried over to the two-sphere.

GOUCHER COLLEGE.

REFERENCES.

[1] R. D. Anderson, "Continuous collections of continuous curves in the plane," *Proceedings of the American Mathematical Society*, vol. 3 (1952), pp. 647-657.

[2] Herman Cohen, "Some results concerning homogeneous plane continua," *Duke Mathematical Journal*, vol. 18 (1951), pp. 467-474.

[3] Mary-Elizabeth Hamstrom, "Concerning continuous collections of continuous curves," *Proceedings of the American Mathematical Society*, vol. 4 (1953), pp. 240-243.

[4] ———, "Concerning webs in the plane," *Transactions of the American Mathematical Society*, vol. 74 (1953), pp. 500-513.

[5] R. D. Anderson and Mary-Elizabeth Hamstrom, "A note on continuous collections of continuous curves filling up a continuous curve in the plane," *Proceedings of the American Mathematical Society*, vol. 5 (1954), pp. 748-752.

[6] Mary-Elizabeth Hamstrom, "Concerning the imbedding of upper semicontinuous collections of continua in continuous collections of continua," *American Journal of Mathematics*, vol. 76 (1954), pp. 793-810.

[7] R. L. Moore and J. R. Kline, "On the most general plane closed point set through which it is possible to pass a simple continuous arc," *Annals of Mathematics*, vol. 20 (1919), pp. 219-223.

[8] R. L. Moore, "Foundations of point set theory," *American Mathematical Society Colloquium Publications*, vol. 13 (1932).

[9] J. H. Roberts, "Collections filling a plane," *Duke Mathematical Journal*, vol. 2 (1936), pp. 10-19.

COMMUTATOR SUBGROUPS OF FREE GROUPS.*[1]

By MAURICE AUSLANDER and R. C. LYNDON.

Let F be a non-abelian free group, R a non-trivial normal subgroup of F, and $G = F/R$. Then F and R are free groups of ranks at least two. Let $[R, R]$ be the commutator subgroup of R, and let $F_0 = F/[R, R]$ and $R_0 = R/[R, R]$. Then R_0 is a free abelian group of rank at least two. Moreover, the inner automorphisms of F induce a natural operation of G on R_0 which we denote by $g \cdot r_0$.

THEOREM 1. *The operation of G on R_0 is effective; that is, only the unit element of G leaves all elements of R_0 fixed.*

Proof. We may suppose $G \neq 1$. Let $g \neq 1$ in G generate a cyclic group G', and let F' be the inverse image of G' under the canonical map of F onto G. Then $G' = F'/R$, and the operation of G' on R_0 agrees with that of G on R_0. We must show that g, and hence that G', does not operate trivially on R_0.

Since $G' = F'/R$ where F' is a free group, the cup product reduction theorem [2, Theorem 10.1] asserts that for any G'-module A

$$H^3(G', A) \simeq H^1(G', \mathrm{Hom}(R_0, A)), \tag{1}$$

where the operation of G' on $\mathrm{Hom}(R_0, A)$ is defined by

$$(g \cdot \theta)(r_0) = g \cdot [\theta(g^{-1} \cdot r_0)], \quad \theta \in \mathrm{Hom}((R_0, A), \quad r_0 \in R_0.$$

We take A to be isomorphic with the cyclic group $G' = Z_n$, $n \geqq 0$, with G' operating trivially on A. Suppose that G' operates trivially on R_0 as well. Then all operations in (1) are trivial. Since G' is cyclic, $H^3(G', A)$ is cyclic or trivial [2, 16.2]. On the other hand $H^1(G', \mathrm{Hom}(R_0, A))$ is isomorphic to $\mathrm{Hom}(Z_n, \mathrm{Hom}(R_0, Z_n))$ and hence to $\mathrm{Hom}(Z_n \otimes R_0, Z_n)$ where $Z_n \otimes R_0$ denotes the tensor product of Z_n and R_0. Since R_0 is a free abelian group of rank at least two, $Z_n \otimes R_0$ is a direct product of at least two copies of Z_n.

---

* Received February 25, 1955; revised May 16, 1955.

[1] Part of the work presented in this paper was done while the first named author was at the University of Chicago. Substantial simplification of the proofs, and a strengthening of Theorem 3, are due to R. L. Taylor.

But $Z_n \neq 1$. Therefore, $\operatorname{Hom}(Z_n \otimes R_0, Z_n)$ is neither cyclic nor trivial. This contradicts (1). Therefore G' does not operate trivially on R_0.

For the following corollaries, we asume that $R \neq F$.

COROLLARY 1.1. *$[R,R]$ is a proper subgroup of $[F,R]$.*

Proof. Let $f \in F$, $f \notin R$. If $[f,r] = frf^{-1}r^{-1}$ were in $[R,R]$ for all r in R, then, setting $g = fR$ in G, we should have $g \cdot r_0 = r_0$ for all $r_0 = r[R,R]$ in R_0. This contradicts Theorem 1.

COROLLARY 1.2. *$[R,R]$ is a proper subgroup of $[F,F]$.*

COROLLARY 1.3. *The center C_0 of F_0 is a proper subgroup of R_0.*

Proof. We first show that $C_0 \subset R_0$. Suppose f_0 is in C_0. Then $f_0 r_0 = r_0 f_0$ for all r_0 in R_0. This implies that the image $g = f_0 R$ of f_0 in G operates trivially on R_0. By Theorem 1, this implies that $g = 1$. Hence f_0 is in R_0.

Suppose that $C_0 = R_0$. Then $f_0 r_0 = r_0 f_0$ for all f_0 in F_0 and r_0 in R_0. Hence all g in G operate trivially on R_0. By Theorem 1, this is possible only if $G = 1$, which contradicts the hypothesis that $R \neq F$.

COROLLARY 1.4. *The center C_0 of F_0 consists precisely of those elements of R_0 that are left fixed under the operation of G.*

THEOREM 2. *The center C_0 of F_0 is trivial if and only if G is infinite.*

Proof. Let G be of finite order n, with elements $g_1 = f_1 R, \cdots, g_n = f_n R$. Suppose $C_0 = 1$. For any r in R, and corresponding $r_0 = r[R,R]$ in R_0, the element $\prod_{i=1}^{n} g_i \cdot r_0$ is fixed under G, hence in C_0. Therefore $\prod g_i \cdot r_0 = 1$. This implies that $\prod f_i r f_i^{-1}[R,R] \subset [F,F]$. Since $f_i r f_i^{-1} \equiv r$ modulo $[F,F]$, this implies that $r^n \in [F,F]$, where $n \geqq 1$. Since $F/[F,F]$ is torsion free, r is in $[F,F]$. Thus $C_0 = 1$ implies $R \subset [F,F]$. But this would imply that the free abelian group $F/[F,F]$ were a homomorphic image of the finite group $G = F/R$, a contradiction. Therefore $C_0 \neq 1$.

Let G be infinite, with elements $g_0 = 1, g_1, g_2, \cdots$. Let ZG be the integer group ring of G. Its elements are of the form $\gamma = \sum n_i g_i$, where only a finite number of the integers n_i are not zero. Let G operate on ZG by left multiplication. If γ were fixed under G, then $g_i^{-1} \cdot \gamma = \gamma$ for all i. This implies that all $n_i = n_0$, and hence that all n_i are zero. Then $\gamma = 0$. It follows that in any free G-module only the zero element is fixed under G. Since [cf. 1, Ch. XIV, Ex. 11] R_0 is a submodule of a free G-module, R_0 is without non-trivial fixed elements. Therefore, by Corollary 1.4, C_0 is trivial.

For any group G, let $G_2 = [G, G]$, $G_3 = [G_2, G]$.

LEMMA. *If F is a non-abelian free group, and R a normal subgroup such that $F_2 \subset R$ and $F_3 \subset R_2$, then $F = R$.*

Proof. From $F_3 \subset R_2$ it follows that $F_2/R_2 \subset C_0$. *If* $R \neq F$, then $R_2 \neq F_2$ by Corollary 1.2, whence $C_0 \neq 1$, and, by Theorem 2, $G = F/R$ is finite. Thus, in any case, G is finite.

The commutator operation determines a homomorphism

$$\kappa\colon\ F/F_2 \otimes F/F_2 \to F_2/F_3.$$

This induces a homomorphism $\mu\colon G \otimes R_0 \to R_0$; explicitly, $\mu(g \circ r_0) = g \cdot r_0 - r_0$. Since G is finite, $G \otimes R_0$ is a torsion group. But R_0 is a free abelian group. It follows that $\mu = 0$. This means that $R_0 \subset C_0$. By Corollary 1.3, this implies that $R = F$.

THEOREM 3. *If S and R are normal subgroups of a non-abelian free group, and $[S, S] \subset [R, R]$, then $S \subset R$.*

Proof. Let $F = RS$. We may suppose that F is a non-abelian free group. We have

$$F_2 = [F, F] = [RS, RS] = [R, R][R, S][S, S] \subset R_2(R \cap S),$$

$$F_3 = [F_2, F] \subset [R_2(R \cap S), RS] \subset R_2[R, R][S, S] = R_2.$$

The lemma now gives $R = F$, whence $S \subset R$.

UNIVERSITY OF MICHIGAN.

REFERENCES.

[1] H. Cartan and S. Eilenberg, Homological Algebra, Princeton University Press, 1955.

[2] S. Eilenberg and S. MacLane, " Cohomology theory in abstract groups. I," *Annals of Mathematics*, vol. 48 (1947), pp. 51-78.

ERRATA.

P. Hartman and A. Wintner, "Asymptotic integrations of linear differential equations," this JOURNAL, vol. 77 (1955), pp. 45-86 and 404.

On page 47, formula (xii), read $h_* - h_0$ in place of $h_0 - h_*$ in the exponent of t.

On page 60, line 13, read (13) in place of (12).

On page 62, formula (67), and on page 72, formula (119), read $\epsilon(\alpha)$ in place of the factor $(\epsilon(j) + \epsilon(\alpha) - 1)$ in the exponent of t.

On page 63, formula (70), read $>$ instead of $<$ (twice); in place of the two lines following (70), read "In this case, $y(t)$ is unique."

On page 64, line 11, read [6] in place of [5].

On page 64, third line of Section 12, read $y \rightarrow e^{-\mu t} y$ in place of $y \rightarrow e^{-\mu t}$.

On page 66, formula (80), read $(b(q) - k)$ in place of the factor $(b(q) - h)$.

On page 67, at the end of formula (87), read $+ k_0$ in place of $- k_0$.

On page 68, formula (91), read $(\epsilon(\alpha) - \epsilon(j))$ in place of the factor $(\epsilon(\alpha) - 1)$ in the exponent of t.

On page 69, in the first line of formula (97), read $k > b$ in place of $k < b$.

On page 83, formula (171), read 1 in place of -1.

On page 83, formulae (172) and (174), read $(1 - f'^2/16f^3)$ in place of $(1 + f'^2/16f^3)$.

On page 84, formula (175), read $(1 - a)$ in place of $(1 + a)$.

On page 84, formula (177), read $(f^{\frac{1}{2}}(t) - a(t))$ in place of $(f^{\frac{1}{2}}(t) + a(t))$.

On page 84, beginning of formula (182), read $(f^{-\frac{1}{2}} z')'$ in place of $(f^{-\frac{1}{2}} z')$.